

AD 743607

THE UNIVERSITY
OF WISCONSIN



MATHEMATICS RESEARCH CENTER

Reproduced by
NATIONAL TECHNICAL
INFORMATION SERVICE
Springfield, Va 22151

Address:

Mathematics Research Center
University of Wisconsin-Madison
Madison, Wisconsin 53706
U.S.A.

150

AR 70-31 **Unclassified**

~~Security Classification~~

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing classification must be entered when the overall report is classified)

1. ORIGINATING AGENCY (Corporate author) Mathematics Research Center, University of Wisconsin, Madison, Wis. 53706		2a. REPORT SECURITY CLASSIFICATION Unclassified	
		2b. GROUP None	
3. REPORT TITLE MINIMAX PROBLEMS, SADDLE FUNCTIONS AND DUALITY			
4. DESCRIPTIVE NOTES (Type of report and inclusive dates) Summary Report; no specific reporting period.			
5. AUTHOR (Last name, first name, middle name) Lynn McLinden			
6. REPORT DATE March 1972		7a. TOTAL NO. OF PAGES 146	7b. NO. OF PAGES 63
8a. CONTRACT OR GRANT NO. Contract No. DA-31-124-ARO-D-462		8b. ORIGINATOR'S REPORT NUMBER #1190	
9. PROJECT NO. None		10. OTHER REPORT NUMBERS (Any other numbers that may be assigned this report) None	
11. DISTRIBUTION STATEMENT Distribution of this document is unlimited.			
12. SUPPLEMENTARY NOTES None		13. ORIGINATOR'S MILITARY ACTIVITY Army Research Office-Durham, N.C.	
14. ABSTRACT Dual pairs of operations for equivalence classes of concave-convex functions are studied. Applications include a perturbational duality theory for minimax problems. A good Lagrange multiplier principle for minimax problems in general is shown to be impossible, and a minimax version of Fenchel's Duality Theorem is proved.			

DD FORM 1473

Unclassified

~~Security Classification~~

400 0100A

**THE UNIVERSITY OF WISCONSIN
MATHEMATICS RESEARCH CENTER**

Contract No. : DA-31-124-ARO-D-462

MINIMAX PROBLEMS, SADDLE FUNCTIONS

AND DUALITY

Lynn McLinden

**This document has been approved for public
release and sale; its distribution is unlimited.**

**MRC Technical Summary Report # 1190
March 1972**

Received November 4, 1971

Madison, Wisconsin 53706

ABSTRACT

In this paper the theory of minimax problems is developed further via the dual approach, that is, by means of the conjugacy correspondence among saddle functions. The saddle functions considered are extended-real-valued, concave in one argument and convex in the other argument. The results obtained extend to minimax problems many of the results already known for convex optimization problems. The proofs, however, are not routine extensions of the ones in the convex case. This is because each minimax problem corresponds in a natural way to a whole equivalence class of saddle functions, and consequently one must always deal with these equivalence classes rather than just with individual functions. In the first half of the paper various operations are described for forming new equivalence classes from given ones. It is shown that these operations fall into dual pairs with respect to the conjugacy correspondence. Included are the important operations of addition and its dual, minimax convolution. Formulas are given describing the effects of the operations on the subdifferential mappings of the equivalence classes. In the second half of the paper, generalized saddle programs are defined and the earlier results are used to develop a perturbational duality theory for such programs. Several characterizations are given for stable optimal solutions and Kuhn-Tucker vectors, including a Lagrangian saddle point characterization. Two special types of programs are then considered. The results for the first type show, somewhat surprisingly, that in general

there does not exist a good Lagrange multiplier principle for minimax problems subject to convex inequality constraints. The results for programs of the second type constitute a minimax version of Fenchel's Duality Theorem. The appendix discusses polyhedral refinements, which are possible for nearly all of the results.

CONTENTS

Introduction	1
A Note to the Reader	14
§0: Preliminaries	15
§1: Two Dual Operations	28
§2: Sharper Results	39
§3: Addition and Minimax Convolution	48
§4: The Partial Conjugacy Operation	68
§5: Generalized Saddle Programs	75
§6: Ordinary Saddle Programs and Lagrange Multipliers	101
§7: Saddle Programs of Fenchel Type	124
Appendix: Polyhedral Refinements	132
References	140

Introduction

In recent years much work has been done on convex optimization problems, and especially on convex programming problems. The dual approach to these problems, which involves applying the theory of conjugate convex functions, has been very successful. It has led to definitive results for convex optimization problems.

In passing to the theory of minimax problems, one encounters a formidable technical difficulty not found in purely convex problems. It is that each minimax problem corresponds not just to a single saddle function but rather to a whole equivalence class of saddle functions. However, a conjugacy correspondence among such equivalence classes has been developed. By means of it the basic questions concerning the existence and nature of saddle points have been fairly well answered, and in the past several years a number of methods have been presented for actually locating saddle points.

The present paper aims to develop further the theory of minimax problems along the lines of the recent results for convex problems. It is hoped that this will serve to complement current efforts toward methods of finding saddle points and to give further impetus to these efforts. Also, this paper is presented in support of the thesis that nearly everything that can be proved for convex optimization problems via the dual approach can similarly be established for minimax problems.

Sponsored by the United States Army under Contract No.: DA-31-124-ARO-D-462 and the United States Air Force under Contract No. AF-AFOSR-71-1994.

Our plan in this introduction is first to review some of the literature pertaining to minimax problems, next to sketch the results for convex optimization problems which this paper extends, then to review very briefly a few basic notions concerning saddle functions, and finally to outline the results obtained.

Minimax theory originated in 1928 with von Neumann's minimax theorem for matrix games [38]. Various proofs and generalizations of this theorem have been given by many authors, including Ville [62], Kakutani [29], Wald [63], Shiffman [53], Fan [20, 21], Kneser [30], Glicksberg [26], Nikaidô [39], Berge [4], Sion [54], Ghoulla-Houri [25], Moreau [37], and Rockafellar [43, 44]. Much of the early work in minimax theory was done in connection with game theory. However in about 1950 two equivalences were established which made it apparent that minimax theory was intimately related to mathematical programming. One of these two equivalences was that between matrix games and dual pairs of linear programs (see Dantzig [13], Gale-Kuhn-Tucker [24], and Charnes [8]). The other equivalence was that between convex programs and Lagrangian saddle point problems (see Kuhn-Tucker [32], Slater [55], and extensions given by Hurwicz-Uzawa in [2]). Various authors, including Stoer [56, 57], Mangasarian-Ponstein [35], and Dantzig-Eisenberg-Cottle [14], later derived duality results for constrained maximization and minimization problems by means of minimax theorems.

In 1964 Rockafellar [43] defined a conjugacy correspondence among saddle functions parallel to that of Fenchel [22] for convex functions. This

correspondence was used in [47] to represent (in finitely many different ways) a certain dual pair of convex programs as a dual pair of minimax problems. At a later date Tynjanskii [60] independently defined the conjugacy correspondence for a more restrictive class of saddle functions. He used it to associate with a given concave-convex game another game of the same type, and showed how solving such a pair of "dual games" is equivalent to solving a related pair of convex programs. Also, papers of Moreau [37] and Ioffe-Tikhomirov [28] contain implicit results concerning the conjugacy correspondence among saddle functions.

The relevance of minimax theory to mathematical economics has long been recognized, dating back to the beginnings of game theory. More recently, minimax theory has been useful in the calculus of variations and optimal control theory (e.g. Rockafellar [50, 51]). It also plays a role in differential games (e.g. [31]).

Related to minimax problems are max-min problems, i.e. two-stage problems of the form $\max_x (\min_y f(x, y))$. These have been studied by Pshenichnyi [40], Danskin [11], and Bram [5]. Such problems correspond to "half" a saddle point problem and arise from such practical considerations as two-stage resource allocation.

The preceding references deal primarily with theory. However the task of actually finding saddle points has also been studied. Work in the early 1950's was done by Brown-von Neumann [6], Robinson [41], and Danskin [10]. Charnes [8] showed that a minimax problem corresponding

to a constrained matrix game is equivalent to a dual pair of linear programs, so that such techniques as the simplex method could be applied. Conversely, in order to utilize the Kuhn-Tucker theorem [32] and its generalizations for solving concave programs, Arrow-Hurwicz [2, p. 118] developed a "steepest descent" method for locating the saddle points of the Lagrangian. Further generalizations of the method of "steepest descent" in connection with saddle points are discussed in Rockafellar [52]. Methods have also been given recently by Demyanov [15, 17, 18], Auslender [3], Danskin [12], Cherruault-Loridan [9], Gratchev-Evtushenko [27], and Čaikovskij [7]. See also Trémolières' survey paper [59]. Methods dealing with max-min problems have been given by Pshenichnyi [40], Demyanov [16], and Danskin [12].

The problem of minimizing a convex function subject to constraints has been analyzed by various authors by means of the duality theory arising from Fenchel's conjugacy correspondence. This dual approach, as expounded in [48], rests ultimately on the duality between two operations which combine a convex function with a linear transformation. In this paper we analyze constrained minimax problems in a similar fashion by means of the duality theory arising from the conjugacy correspondence among saddle functions. To accomplish this we develop for saddle functions analogues of these fundamental operations on convex functions. But before actually describing our results, we shall sketch the two operations and the applications of them which this paper extends.

One of the two operations is just to form the composition fA of a convex function f with a linear transformation A . The other operation may be called taking the image of f under A , and the resulting function Af is defined by $(Af)(x) = \inf\{f(y) | Ay = x\}$. The fundamental result connecting these operations is that, under certain mild hypotheses,

$$(fA)^* = A^* f^* , \quad (1)$$

where $*$ of a linear transformation denotes the adjoint linear transformation and $*$ of a convex function denotes the conjugate convex function.

One of the main consequences of the duality formula (1) is the duality between the operations of addition and infimal convolution for convex functions. This can be obtained by taking f to be the separable function

$$f(x_1, \dots, x_m) = f_1(x_1) + \dots + f_m(x_m) ,$$

where each f_i is convex on R^n , and defining A to be the linear transformation which sends each element x of R^n into the m -tuple (x, \dots, x) . In this event fA is $f_1 + \dots + f_m$ and $A^* f^*$ is the function

$$x^* \rightarrow \inf\{f_1^*(x_1^*) + \dots + f_m^*(x_m^*) | x^* = x_1^* + \dots + x_m^*\} ,$$

i.e. the infimal convolution of f_1^*, \dots, f_m^* . Formula (1) then implies that, under mild hypotheses, the conjugate of the sum is the infimal convolute of the conjugates. This gives a framework encompassing problems of the form, "minimize $h(x)$ subject to $x \in C$," where h and C are convex. Simply take $m = 2$, let $f_1 = h$, and let $f_2(x)$ equal 0 when $x \in C$ and $+\infty$ otherwise.

The duality represented by formula (1) is also fundamental in the perturbational duality theory developed by Rockafellar for generalized convex programs [48]. Among other things, this theory generalizes the classical results about dual linear programs and generalizes Fenchel's Duality Theorem [22, p. 108] (see also [45, 46] and Stoer-Witzgall [58]). It also sheds light on the Lagrange multiplier principle for convex programming and thereby on the celebrated Dantzig-Wolfe decomposition principle for linear and convex programs [48, pp. 285-290] (see also Falk [19] and Lasdon [33]).

For convenient reference, in the next three paragraphs we review some basic definitions and facts about saddle functions due to Rockafellar [43].

There is an equivalence relation among saddle functions which has the property that equivalent saddle functions have the same (lower and upper) saddle values and also the same saddle points. The relation is the following: two concave-convex functions K and L are said to be equivalent if and only if for each x the closures of the convex functions $K(x, \cdot)$ and $L(x, \cdot)$ coincide and for each y the closures of the concave functions $K(\cdot, y)$ and $L(\cdot, y)$ coincide. The equivalence classes determined by this equivalence relation are what we take to be the natural objects of study in minimax theory.

Recall that in convex function theory, in order for the crucial duality formula

$$(f^*)^* = f \quad (2)$$

to hold, one considers convex functions which are closed, i. e. lower semi-continuous. This is a natural, constructive regularity assumption to make. Similarly, in saddle function theory an analogue of formula (2) holds for "regularized" saddle functions. A saddle function K is defined to be closed if and only if it is equivalent to both its convex closure and its concave closure, where by convex (resp. concave) closure we mean the saddle function obtained from K by closing it (in the sense of convex function theory) in its convex (resp. concave) argument. Trivially, a saddle function is closed if and only if every member of its equivalence class is closed. Furthermore, the property of being closed is a constructive regularity condition for saddle functions. Equivalent closed saddle functions must be very nearly equal in that they can differ essentially only at the "corner points" of their "domain of finiteness." Moreover, each equivalence class $[K]$ of closed saddle functions is an interval in the sense that there exist unique members \underline{K} and \bar{K} of $[K]$ such that $[K]$ contains all, and only those, saddle functions \tilde{K} satisfying $\underline{K} \leq \tilde{K} \leq \bar{K}$.

If K is a concave-convex function from $R^m \times R^n$ to $[-\infty, +\infty]$, the lower conjugate \underline{K}^* and upper conjugate \bar{K}^* of K are defined by

$$\underline{K}^*(x^*, y^*) = \sup_y \inf_x \{ \langle x, x^* \rangle + \langle y, y^* \rangle - K(x, y) \}$$

and

$$\bar{K}^*(x^*, y^*) = \inf_x \sup_y \{ \langle x, x^* \rangle + \langle y, y^* \rangle - K(x, y) \}.$$

These functions are concave-convex. When K is closed, \underline{K}^* and \bar{K}^* are

equivalent and closed and, moreover, they depend only on the equivalence class $[K]$ containing K . Thus, associated with each equivalence class $[K]$ of closed concave-convex functions is another well-defined equivalence class $[K^*]$ of closed concave-convex functions, namely the class containing \underline{K}^* and \bar{K}^* . The class $[K^*]$ is said to be the conjugate of $[K]$. This conjugacy correspondence has the crucial property that the conjugate of $[K^*]$ is $[K]$, which is the analogue of formula (2) for saddle functions.

With this review of general facts in mind, we now describe the results obtained in this paper. We begin with the analogues of the two fundamental operations described above. Suppose K is a closed concave-convex function and that A is the linear transformation $A_1 \times A_2$ obtained from two other linear transformations A_1 and A_2 by $A_1 \times A_2(x, y) = (A_1 x, A_2 y)$. One of our operations consists of forming an equivalence class $[KA]$ containing all saddle functions of the form

$$(x, y) \rightarrow \tilde{K} A(x, y) = \tilde{K}(A_1 x, A_2 y)$$

for \tilde{K} any member of $[K]$. A mild hypothesis is given which ensures that in fact such a single class exists and, moreover, that all its members are closed. The other operation is to form a single equivalence class $[AK]$ containing all saddle functions both of the form

$$(u, v) \rightarrow \sup_{\{x | A_1 x = u\}} \inf_{\{y | A_2 y = v\}} \tilde{K}(x, y)$$

and of the form

$$(u, v) \rightarrow \inf_{\{y | A_2 y = v\}} \sup_{\{x | A_1 x = u\}} \tilde{K}(x, y)$$

for \tilde{K} any member of $[K]$. A mild hypothesis is given which ensures that indeed such a class exists and that all of its members are closed. What is surprising is that this hypothesis is precisely the same as is needed to ensure the existence of the class $[K^* A^*]$ formed by the first operation from $[K^*]$ and $A^* = A_1^* \times A_2^*$. Furthermore, it is shown that under this hypothesis $[AK]$ and $[K^* A^*]$ are actually conjugate classes. This is the analogue of formula (1) for saddle functions.

The development of these operations and the proof of the duality between them form the heart of this paper. Two forms of this duality are given. The form given in §1 is the more widely and easily applicable. But the sharper form in §2 gives especially strong conclusions concerning the nature of the equivalence class $[AK]$, including information about the attainment of the minimax extrema appearing in its definition.

In §3 the first application of this duality is made in defining addition and minimax convolution for saddle functions and showing that these are dual operations. The formula

$$\partial(K_1 + K_2)(x, y) = \partial K_1(x, y) + \partial K_2(x, y)$$

is obtained for the subdifferential of the sum of two saddle functions. This parallels the result for convex functions obtained by Rockafellar [42], Moreau [36], and others. The duality between addition and minimax convolution gives a general framework within which to consider problems of the form, "find the saddle points of H with respect to $C \times D$," where H is a saddle function and C and D are convex sets.

From the results on addition and minimax convolution we obtain an interesting byproduct. For $i = 1, \dots, p$ let K_i be a closed concave-convex function on $\mathbb{R}^m \times \mathbb{R}^n$ which is not identically $+\infty$ or $-\infty$ and let T_i be the maximal monotone operator on \mathbb{R}^{m+n} arising from the subdifferential of K_i (see [48,49]). If each $R(T_i)$ is bounded, where $R(\cdot)$ denotes the range of an operator, then ΣT_i is maximal monotone and

$$\Sigma \text{cl } R(T_i) = \text{cl } R(\Sigma T_i) . \quad (3)$$

It is known that this formula holds whenever the T_i 's are subdifferentials of closed proper convex functions and each $R(T_i)$ is bounded. On the other hand, formula (3) fails in general for maximal monotone operators. However it is not known whether formula (3) holds for arbitrary maximal monotone operators under the assumption that the sets $R(T_i)$ are bounded. But the fact that it holds for those maximal monotone operators arising from saddle functions leads one to conjecture it holds in general. This is because such operators, unlike the subdifferentials of convex functions, exhibit most of the pathology of arbitrary maximal monotone operators. Indeed, this last fact is one of the main motivations for studying saddle functions.

In §5, as a second principal application of our fundamental dual operations, we develop a perturbational duality theory for generalized saddle programs. We define a generalized saddle program to be an "objective" saddle function K_0 (thought of as some given minimax problem) together with a particular class of perturbations. The entire program is given by another

saddle function K . To this generalized saddle program K we associate a dual generalized saddle program L . Under mild hypotheses on the perturbations in K , the dual program L has a unique (up to equivalence) "objective" saddle function L_0 . The minimax problem corresponding to L_0 is a dual to the original minimax problem. One of the main features of this theory is that choosing different classes of perturbations of the original problem gives rise to different dual problems. Optimal solutions, stable optimal solutions and Kuhn-Tucker vectors for these dual saddle programs are studied and various duality theorems are proved. In §4, as a subsidiary application of the dual operations of §1, a symmetric one-to-one "partial conjugacy" correspondence is defined among equivalence classes of closed saddle functions. By means of this new correspondence we are able to associate with a generalized saddle program and its dual a well-defined Lagrangian saddle function. We then give a characterization of the primal and dual stable optimal solutions and Kuhn-Tucker vectors in terms of the saddle points of the Lagrangian.

In §6 this perturbational duality theory is used to study the problem of finding a saddle point subject to convex and concave inequality constraints. Ordinary saddle programs are defined as a framework to treat such problems. A question of particular concern is whether or not a good Lagrange multiplier principle holds for these saddle programs. The analogous question for ordinary convex programs (i.e. minimizing a convex function subject to convex inequality constraints) has a very satisfying affirmative answer which leads to the important decomposition principle for separable convex programs (see [48, Theorem 28.1 and pp. 285-290]). However we show that such a good Lagrange multiplier principle does not hold in general for ordinary saddle

programs. The reason for this is essentially that, unlike the convex program case, the set of saddle points of the Lagrangian does not split up into the product of the primal stable optimal solutions and the primal Kuhn-Tucker vectors (Lagrange multipliers). Put another way, the stable optimal solutions and Kuhn-Tucker vectors are shown to be in a certain sense dependent on each other. We conclude the section with an explicit description of the dual saddle program.

Finally, in §7 the perturbational duality theory is applied to another class of problems to yield a minimax version of Fenchel's Duality Theorem. We deal with dual pairs of minimax problems of the following form (where for simplicity now we suppress the issue of the domains of the variables):

- (I) Find the saddle points of $K(x, y) - LA(x, y)$,
- (II) Find the saddle points of $L^*(z, w) - K^* A^*(z, w)$.

Here K is closed and concave-convex on $R^m \times R^n$, L is closed and convex-concave on $R^p \times R^q$, and A is a product linear transformation from $R^m \times R^n$ to $R^p \times R^q$. The results obtained generalize certain results of Rockafellar [47], Lebedev-Tynjanskii [34], and Tynjanskii [60, 61].

It is known that many results in the theory of convex functions allow refinements when polyhedrality is present. For closed saddle functions there is a property of polyhedrality which is preserved under conjugacy and also the operations we develop in §§1, 3 and 4. Nearly all the results in the paper admit refinements when such polyhedrality is present. This is discussed in the Appendix.

This paper is the outgrowth of an investigation of a problem posed by Professor R. Tyrrell Rockafellar. The problem was to see if one could develop well-defined ways of forming new equivalence classes of saddle functions from given ones. The author wishes to express his deep gratitude to Professor Rockafellar for posing this problem and for his many helpful comments.

A Note to the Reader

In general, we use numbers enclosed in square brackets to indicate bibliographical references. However a special abbreviation is employed in citing results from Rockafellar [48], due to the frequency with which these are used. Namely, we omit entirely the reference to [48] and merely give in parentheses the number of the result being cited. For example, Theorem 23.8 of [48] is cited simply as (23.8), Corollary 6.3.1 as (6.3.1), and so forth.

Throughout the paper expressions sometimes appear which involve taking the supremum or infimum of an empty set of numbers. Whenever these occur they are to be interpreted using the conventions $\sup \emptyset = -\infty$ and $\inf \emptyset = +\infty$.

§0. Preliminaries

The definitions and notations used in this paper are mainly those set forth in Rockafellar [48]. In this section we review for convenience some of these and also introduce some of our own. In addition, we present a few background results which will be of use later on.

The topology taken on R^n is the usual one, and the interior and closure of a subset S of R^n are denoted by $\text{int } S$ and $\text{cl } S$, respectively. A set is called affine if and only if it is either the empty set, denoted by \emptyset , or a translate of a linear subspace. The affine hull of a subset is the smallest affine set containing it. If C is a convex subset of R^n , its relative interior, written $\text{ri } C$, is the interior of C taken with respect to its affine hull equipped with the relative topology.

If A is a linear transformation from R^p to R^m , then A^* denotes the adjoint linear transformation mapping R^m to R^p .

The effective domain of a convex function f on R^n is the set

$$\text{dom } f = \{x \mid f(x) < +\infty\},$$

and the conjugate of f is the convex function f^* on R^n given by

$$f^*(x^*) = \sup_x \{ \langle x, x^* \rangle - f(x) \}$$

(where $\langle \cdot, \cdot \rangle$ denotes the ordinary inner product). Similarly, the effective domain of a concave function g on R^n is the set

$$\text{dom } g = \{x | g(x) > -\infty\} ,$$

and the conjugate of g is the concave function g^* on R^n given by

$$g^*(x^*) = \inf_x \{ \langle x, x^* \rangle - g(x) \} .$$

Our multiple use of the superscript $*$ should cause no difficulty, since it is always clear from the context what operation is intended.

For any subset C of R^n the function $\delta(\cdot | C)$ on R^n , called the indicator function of C , is defined by setting $\delta(x | C)$ equal to 0 if $x \in C$ and $+\infty$ otherwise. Clearly C is convex if and only if $\delta(\cdot | C)$ is convex, and in this case the conjugate of $\delta(\cdot | C)$ is denoted by $\delta^*(\cdot | C)$ and is given by

$$\delta^*(x^* | C) = \sup \{ \langle x, x^* \rangle | x \in C \} .$$

This function is called the support function of C .

A concave-convex function on $R^m \times R^n$ is a function K from $R^m \times R^n$ to $[-\infty, +\infty]$ such that $K(x, y)$ is a concave function of $x \in R^m$ for each fixed $y \in R^n$ and a convex function of $y \in R^n$ for each fixed $x \in R^m$. A convex-concave function is defined the same except for interchanging "concave" with "convex." A saddle function is either a concave-convex or a convex-concave function.

For the remainder of §0 let K denote a concave-convex function on $R^m \times R^n$. For convex-concave functions we make the obvious changes in the definitions which follow.

We say that K has a saddle value, or that the saddle value exists, if and only if the two quantities

$$\sup_x \inf_y K(x, y)$$

and

$$\inf_y \sup_x K(x, y)$$

are equal, in which case this common value is the saddle value of K . A pair $(\bar{x}, \bar{y}) \in R^m \times R^n$ is a saddle point of K if and only if

$$K(x, \bar{y}) \leq K(\bar{x}, \bar{y}) \leq K(\bar{x}, y)$$

for each $(x, y) \in R^m \times R^n$.

Define subsets $\text{dom}_1 K$ of R^m and $\text{dom}_2 K$ of R^n by

$$\text{dom}_1 K = \{x \mid K(x, \cdot) \text{ is never } -\infty\},$$

$$\text{dom}_2 K = \{y \mid K(\cdot, y) \text{ is never } +\infty\}.$$

The product set

$$\text{dom}_1 K \times \text{dom}_2 K = \text{dom } K$$

is the effective domain of K . We say that K is proper if and only if its effective domain is nonempty. The kernel of K is the restriction of K to the relative interior of its effective domain. We say that K is simple if and only if $\text{dom } K(x, \cdot) \subset \text{cl}(\text{dom}_2 K)$ for every $x \in \text{ri}(\text{dom}_1 K)$ and $\text{dom } K(\cdot, y) \subset \text{cl}(\text{dom}_1 K)$ for every $y \in \text{ri}(\text{dom}_2 K)$. The function $\text{cl}_1 K$ obtained by closing $K(x, y)$ as a concave function of x for each fixed y is called the concave closure of K . Similarly, the function $\text{cl}_2 K$ obtained by closing $K(x, y)$ as a convex function of

y for each fixed x is called the convex closure of K . If L is also a concave-convex function on $R^m \times R^n$, we say that K and L are equivalent and write $K \sim L$ if and only if $cl_1 K = cl_1 L$ and $cl_2 K = cl_2 L$. The collection of all concave-convex functions on $R^m \times R^n$ which are equivalent to K is called the equivalence class containing K and is denoted by $[K]$. We say that K is closed if and only if $cl_1 K \sim K$ and $K \sim cl_2 K$.

It is an easy exercise to show that $K \sim L$ implies both that $\text{dom } K = \text{dom } L$ and that K is closed if and only if L is closed. Thus, an equivalence class is called closed (resp. proper) if and only if any of its members is closed (resp. proper).

The function f on $R^m \times R^n$ given by $f(x, y^*) = \sup_y \{ \langle y, y^* \rangle - K(x, y) \}$ is convex in (x, y^*) jointly, since it is a pointwise supremum of convex functions. Similarly, the function g on $R^m \times R^n$ given by $g(x^*, y) = \inf_x \{ \langle x, x^* \rangle - K(x, y) \}$ is concave in (x^*, y) jointly. We call f (resp. g) the convex (resp. concave) parent of K . Notice that this usage of the term "parent" differs by some minus signs from the original usage in Rockafellar [47]. With these definitions, (34.2) implies the following.

THEOREM 0.1. Let f (resp. g) be the convex (resp. concave) parent of K . Then K is closed if and only if $f(x, y^*) = -g^*(x, -y^*)$ and $g(x^*, y) = -f^*(-x^*, y)$, in which case (a) and (b) below hold.

(a) The parents of K depend only on $[K]$; that is, for each $\tilde{K} \in [K]$, the convex (resp. concave) parent of \tilde{K} is f (resp. g). Moreover, f and g are closed, and

$$\text{dom}_1 K = \{x \mid f(x, y^*) < +\infty \text{ for some } y^*\},$$

$$\text{dom}_2 K = \{y \mid g(x^*, y) > -\infty \text{ for some } x^*\}.$$

(b) The equivalence class $[K]$ consists of all and only those concave-
convex functions \tilde{K} on $R^m \times R^n$ which satisfy $K \leq \tilde{K} \leq \bar{K}$, where

$$\underline{K}(x, y) = \sup_{y^*} \{ \langle y^*, y \rangle - f(x, y^*) \}$$

and

$$\bar{K}(x, y) = \inf_{x^*} \{ \langle x^*, x \rangle - g(x^*, y) \}.$$

Moreover, $\text{cl}_2 \tilde{K} = \underline{K}$ and $\text{cl}_1 \tilde{K} = \bar{K}$ for each $\tilde{K} \in [K]$, and

$$\underline{K}(x, y) = \bar{K}(x, y)$$

whenever $x \in \text{ri}(\text{dom}_1 K)$ or $y \in \text{ri}(\text{dom}_2 K)$.

The lower conjugate of K , denoted by \underline{K}^* , is the function on $R^m \times R^n$ defined by

$$\underline{K}^*(x^*, y^*) = \sup_y \inf_x \{ \langle x, x^* \rangle + \langle y, y^* \rangle - K(x, y) \}.$$

Similarly, the upper conjugate of K , denoted by \bar{K}^* , is the function on $R^m \times R^n$ defined by

$$\bar{K}^*(x^*, y^*) = \inf_x \sup_y \{ \langle x, x^* \rangle + \langle y, y^* \rangle - K(x, y) \}.$$

From (37.1) we have the following result.

THEOREM 0.2. Assume K is closed. Then \bar{K}^* and \underline{K}^* are equivalent, closed concave-convex functions which depend only on $[K]$. Moreover, if L is any element of the equivalence class containing \bar{K}^* and \underline{K}^* , then

$$cl_2 L = \underline{K}^*, \quad cl_1 L = \bar{K}^*,$$

$$\underline{L}^* = \underline{K}, \quad \bar{L}^* = \bar{K},$$

and the convex (resp. concave) parent of L is the negative of the concave (resp. convex) parent of K .

When it exists, the equivalence class containing \underline{K}^* and \bar{K}^* is called the conjugate of $[K]$ and is denoted by $[K^*]$. Each member of $[K^*]$ is called a conjugate of every member of $[K]$. It is immediate from Theorems 0.2 and 0.1(b) that (at least when K is closed) \underline{K}^* and \bar{K}^* are the least and greatest elements of $[K^*]$, respectively. The notation thus conforms to that introduced in Theorem 0.1(b), where a lower (resp. upper) bar indicates the least (resp. greatest) element of the equivalence class.

By (34.2.3) the only equivalence classes which are closed but not proper are the one containing the constant function $+\infty$ and the one containing the constant function $-\infty$. Since each of these two equivalence classes is the conjugate of the other, it follows that $[K^*]$ is closed and proper whenever $[K]$ is.

We define K to be polyhedral if and only if it is closed and either its concave or its convex parent is polyhedral. If K is polyhedral and L is equivalent to K , then Theorem 0.1 implies L is polyhedral. Thus, an

equivalence class is called polyhedral if and only if any of its members is polyhedral. From Theorem 0.2 follows the important fact that $[K^*]$ is polyhedral whenever $[K]$ is polyhedral. It will be shown that polyhedrality is preserved also by each of the operations developed in §§1, 3 and 4.

The notions of "recession function" and "recession cone" for saddle functions will be quite useful for formulating various growth conditions needed later on. Recall that if f is a proper convex function on R^n , the recession function of f , written $\text{rec } f$, is the function on R^n defined by

$$(\text{rec } f)(y) = \sup\{f(x+y) - f(x) \mid x \in \text{dom } f\},$$

and the recession cone of f is the set

$$\text{rec cone } f = \{y \mid (\text{rec } f)(y) \leq 0\}.$$

The recession function and recession cone of a proper concave function are defined similarly by replacing "sup" by "inf" and " \leq " by " \geq ". This notation for these objects differs from that in [48]. Now write $C = \text{dom}_1 K$ and $D = \text{dom}_2 K$. The convex recession function of K is the function $\text{rec}_2 K$ on R^n defined by

$$(\text{rec}_2 K)(w) = \sup\{(\text{rec } K(x, \cdot))(w) \mid x \in \text{ri } C\}.$$

The convex recession cone of K is the set

$$\text{rec cone}_2 K = \{w \mid (\text{rec}_2 K)(w) \leq 0\}.$$

Similarly, the concave recession function of K is the function $\text{rec}_1 K$ on R^m defined by

$$(\text{rec}_1 K)(z) = \inf\{(\text{rec } K(\cdot, y))(z) \mid y \in \text{ri } D\},$$

and the concave recession cone of K is the set

$$\text{rec cone}_1 K = \{z \mid (\text{rec}_1 K)(z) \geq 0\} .$$

Trivially,

$$\text{rec cone}_2 K = \bigcap \{\text{rec cone } K(x, \cdot) \mid x \in \text{ri } C\}$$

and

$$\text{rec cone}_1 K = \bigcap \{\text{rec cone } K(\cdot, y) \mid y \in \text{ri } D\} .$$

The main importance of the recession functions of K rests on the following theorem, which is (37.2) reformulated.

THEOREM 0.3. Assume K is closed and proper. Then

$$\text{rec}_2 K = \delta^*(\cdot \mid \text{dom}_2 K^*)$$

and

$$\text{rec}_1 K = -\delta^*(-\cdot \mid \text{dom}_1 K^*) .$$

This says that the recession functions of K are essentially just the support functions of $\text{dom } K^*$. But K^* , and hence $\text{dom } K^*$, depends only on $[K]$. Thus, for K closed and proper, the recession functions and recession cones of K depend only on $[K]$.

Suppose now that K is closed and proper. We know by (37.1.3) that the saddle value of K exists if either $0 \in \text{ri}(\text{dom}_1 K^*)$ or $0 \in \text{ri}(\text{dom}_2 K^*)$, and (37.5.3) tells us that a saddle point of K exists if actually $(0, 0) \in \text{ri}(\text{dom } K^*)$. In addition, it is not hard to show that the set of saddle points of K is nonempty and bounded if and only if $(0, 0) \in \text{int}(\text{dom } K^*)$. The next three lemmas will help us to utilize these basic facts.

LEMMA 0.4. Assume K is closed, and let f (resp. g) be its convex
(resp. concave) parent. Then

$$\text{ri}(\text{dom}_1 K^*) = \bigcup \{ \text{ri}(\text{dom } g(\cdot, y)) \mid y \in \text{ri } D \}$$

and

$$\text{ri}(\text{dom}_2 K^*) = \bigcup \{ \text{ri}(\text{dom } f(x, \cdot)) \mid x \in \text{ri } C \} ,$$

where $C = \text{dom}_1 K$ and $D = \text{dom}_2 K$. These formulas also hold when "ri" is
deleted throughout.

PROOF. From Theorems 0.2 and 0.1(a) it follows that

$$\text{dom}_1 K^* = A \text{ dom } g$$

and

$$\text{dom } g = \bigcup \{ \text{dom } g(\cdot, y) \times \{y\} \mid y \in D \} ,$$

where A is the projection $(x^*, y) \rightarrow x^*$. Hence (6.6) implies

$$\text{ri}(\text{dom}_1 K^*) = A \text{ ri}(\text{dom } g)$$

and (6.8) implies

$$\text{ri}(\text{dom } g) = \bigcup \{ \text{ri}(\text{dom } g(\cdot, y)) \times \{y\} \mid y \in \text{ri } D \} .$$

The formulas for $\text{dom}_1 K^*$ and its relative interior follow from these, and the other two formulas are proved similarly.

LEMMA 0.5. Assume K is closed and proper. Let j equal 1 or 2 and
put $S_j = \text{rec cone}_j K$. Then

$$0 \in \text{ri}(\text{dom}_j K^*) \text{ if and only if } S_j \subset -S_j ,$$

and

$0 \in \text{int}(\text{dom}_1 K^\circ)$ if and only if $S_1 \subset (0)$.

PROOF. We use the following SUBLEMMA. If $w^\circ \in R^n$ and h is a positively homogeneous proper convex function on R^n , then the following two conditions are equivalent:

- (a) $\forall w \in R^n, \langle w, w^\circ \rangle \leq h(w)$ with strict inequality for each w such that $-h(-w) \neq h(w)$;
- (b) $\forall w \in R^n, h(w) \leq \langle w, w^\circ \rangle$ implies $h(-w) \leq \langle -w, w^\circ \rangle$.

PROOF OF SUBLEMMA. Assume (a) and suppose $h(w) \leq \langle w, w^\circ \rangle$. Then actually $h(w) = \langle w, w^\circ \rangle$. If we had $-h(-w) \neq h(w)$, then (a) would imply $\langle w, w^\circ \rangle < h(w)$, a contradiction. Thus $-h(-w) = h(w) = \langle w, w^\circ \rangle$ and (b) is proved. Conversely, assume (b) and let w be given. If $h(w) \leq \langle w, w^\circ \rangle$, then (4.7.2) and (b) imply $-h(w) \leq h(-w) \leq \langle -w, w^\circ \rangle$ and hence $\langle w, w^\circ \rangle \leq h(w)$. Suppose $-h(-w) \neq h(w)$. By (4.7.2) we have $-h(-w) < h(w)$. If $h(w) \leq \langle w, w^\circ \rangle$, this would imply $-h(-w) < \langle w, w^\circ \rangle$ while at the same time from (b) we would have $\langle w, w^\circ \rangle \leq -h(-w)$. Therefore $\langle w, w^\circ \rangle < h(w)$ whenever $-h(-w) \neq h(w)$, and (a) is proved.

Now to prove the lemma we define $h = \delta^\circ(\cdot | \text{dom}_2 K^\circ)$. By (13.1) and the Sublemma, $0 \in \text{ri}(\text{dom}_2 K^\circ)$ if and only if for each $w \in R^n$, $h(w) \leq 0$ implies $h(-w) \leq 0$. By (13.1) we also have that $0 \in \text{int}(\text{dom}_2 K^\circ)$ if and only if for each $w \in R^n$, $h(w) \leq 0$ implies $w = 0$. The assertions for $j = 2$ follow from these equivalences and Theorem 0.3. The assertions for $j = 1$ are proved similarly.

Concerning the condition $S_j \subset -S_j$ in Lemma 0.5, recall from (13.2) that the support functions of nonempty convex sets are precisely those

functions whose epigraphs are closed convex cones containing the origin. This fact and Theorem 0.3 imply that the recession cones cl of a closed proper K are themselves closed convex cones containing the origin. Hence they are actually subspaces if and only if they are closed under scalar multiplication by -1 .

The conditions given in the next lemma may well be the easiest to verify in many cases.

LEMMA 0.6. Assume K is closed and proper. Then $(0, 0) \in \text{int}(\text{dom} K^\circ)$ whenever $\text{dom} K$ is bounded. In particular, $0 \in \text{int}(\text{dom}_2 K^\circ)$ whenever there exists an $x \in \text{dom}_1 K$ such that the level sets $\{y \mid K(x, y) \leq \alpha\}$, $\alpha \in \mathbb{R}$, are all bounded, and $0 \in \text{int}(\text{dom}_1 K^\circ)$ whenever there exists a $y \in \text{dom}_2 K$ such that the level sets $\{x \mid K(x, y) \geq \alpha\}$, $\alpha \in \mathbb{R}$, are all bounded.

PROOF. By (34.3), the first assertion is a special case of the second. Now let $x \in \text{dom}_1 K$ be such that all the level sets of $K(x, \cdot)$ are bounded. Then by (7.6), the nonempty level sets of $cl(K(x, \cdot))$ are all bounded. Also, $cl(K(x, \cdot)) = cl_2 K(x, \cdot) = \underline{K}(x, \cdot)$. Hence (8.7) and (8.4) imply that $\text{rec cone } \underline{K}(x, \cdot) = \{0\}$, which by (13.3.4c) is equivalent to $0 \in \text{int}(\text{dom} \underline{K}(x, \cdot)^\circ)$. But $\underline{K}(x, \cdot)^\circ = f(x, \cdot)$, where f is the convex parent of K . Since Lemma 0.4 yields $\text{int}(\text{dom } f(x, \cdot)) \subset \text{int}(\text{dom}_2 K^\circ)$, it follows that $0 \in \text{int}(\text{dom}_2 K^\circ)$. The statement about $0 \in \text{int}(\text{dom}_1 K^\circ)$ is proved similarly.

As explained above, Lemmas 0.4 through 0.6 furnish various criteria for the existence of saddle values and saddle points. To facilitate the use of these criteria, in §§1 and 3 we derive formulas for the parents and

recession functions of the saddle functions which result from various operations. By combining these formulas with the preceding criteria, the reader can derive existence theorems as needed. Consequently, throughout this paper existence results will not be stressed.

We conclude this section with a lemma which will be useful later on in dualizing various conditions with respect to the conjugacy correspondence.

LEMMA 0.7. Assume K is closed and proper, and let L_1 and L_2 be subspaces of R^m and R^n . Then for $j = 1$ and 2 the following conditions are equivalent:

$$(a_j) \quad L_j \cap \text{ri}(\text{dom}_j K^*) \neq \emptyset ;$$

$$(b_j) \quad L_j^\perp \cap (\text{rec cone}_j K) \text{ is a subspace} ;$$

$$(c_j) \quad L_j^\perp \cap (\text{rec cone}_j K) \subset -(\text{rec cone}_j K) .$$

PROOF. We prove only that (a_2) , (b_2) and (c_2) are equivalent, as the proof for $j = 1$ is similar. By the remark following Lemma 0.5 and the fact that L_2 is a subspace, (b_2) is equivalent to (c_2) . Write $D^* = \text{dom}_2 K^*$ and $L = L_2$. By (11.3), (a_2) fails if and only if there exists a hyperplane separating L and D^* properly. By (11.1) this occurs if and only if there exists a $w \in R^n$ such that

$$\inf\{\langle y^*, w \rangle \mid y^* \in L\} \geq \sup\{\langle y^*, w \rangle \mid y^* \in D^*\}$$

and

$$\sup\{\langle y^*, w \rangle \mid y^* \in L\} > \inf\{\langle y^*, w \rangle \mid y^* \in D^*\} .$$

But since

$$\sup\{\langle y^*, w \rangle \mid y^* \in D^*\} = (\text{rec}_2 K)(w)$$

and

$$\inf\{\langle y^*, w \rangle \mid y^* \in L\} = \begin{cases} 0 & \text{if } w \in L^\perp \\ -\infty & \text{if } w \notin L^\perp, \end{cases}$$

this means that (a_2) fails if and only if there exists a $w \in L^\perp$ such that $(\text{rec}_2 K)(w) \leq 0$ and $(\text{rec}_2 K)(-w) > 0$. Therefore (a_2) holds if and only if for each $w \in L^\perp$, $(\text{rec}_2 K)(w) \leq 0$ implies $(\text{rec}_2 K)(-w) \leq 0$. But this last condition is the same as (c_2) .

§1. Two Dual Operations

In this section we develop two quite general ways of forming new equivalence classes of saddle functions from given ones. When viewed as operations on equivalence classes, they are seen to be dual to each other with respect to the basic conjugacy correspondence. Various results are proved concerning the equivalence classes resulting from these operations. The section concludes with examples showing that the conditions under which the operations can be performed cannot in general be weakened.

The first operation we develop is analogous to that of composing a convex function with a linear transformation. Let K be a closed proper concave-convex function on $R^m \times R^n$ and let $A = A_1 \times A_2$ be a linear transformation from $R^p \times R^q$ to $R^m \times R^n$. We seek a condition on K and A ensuring the existence of an equivalence class of closed proper saddle functions which contains every function of the form $\tilde{K}A$ for $\tilde{K} \in [K]$. Such a condition is given in Theorem 1.2. When this equivalence class exists, it will usually be denoted by $[KA]$.

LEMMA 1.1. Let K be a concave-convex function on $R^m \times R^n$ and let $A = A_1 \times A_2$ be a linear transformation from $R^p \times R^q$ to $R^m \times R^n$. Then KA is a concave-convex function on $R^p \times R^q$, and $A^{-1}(\text{dom } K) \subset \text{dom } KA$. The inclusion can be strengthened to equality if K is closed and proper and $\text{range } A \cap \text{ri}(\text{dom } K) \neq \emptyset$.

PROOF. Write $\text{dom } K = C \times D$. By (5.7), KA is concave-convex. If $u \in A_1^{-1}C$, then $K(A_1 u, \cdot)$ is never $-\infty$ and hence $KA(u, \cdot) = K(A_1 u, \cdot)A_2$ is

never $-\infty$. This shows $A_1^{-1}C \subset \text{dom}_1 KA$. Similarly $A_2^{-1}D \subset \text{dom}_2 KA$. Now assume K is closed and proper and $\text{range } A \cap \text{ri}(\text{dom } K) \neq \emptyset$. If $u \notin A^{-1}C$, then (34.3) implies $K(A_1 u, \cdot) A_2$ equals $-\infty$ everywhere on $\text{ri } D$ and hence $K(A_1 u, \cdot) A_2$ equals $-\infty$ everywhere on $A_2^{-1}(\text{ri } D)$. Since $A_2^{-1}(\text{ri } D) \neq \emptyset$ by hypothesis, this shows $\text{dom}_1 KA \subset A_1^{-1}C$. Similarly $\text{dom}_2 KA \subset A_2^{-1}D$.

THEOREM 1.2. Let K be a closed proper concave-convex function on $R^m \times R^n$, and let $A = A_1 \times A_2$ be a linear transformation from $R^p \times R^q$ to $R^m \times R^n$ such that $\text{range } A \cap \text{ri}(\text{dom } K) \neq \emptyset$. Then the collection $\{\underline{KA} | \bar{K} \in [K]\}$ of saddle functions is contained in an equivalence class $[H]$ of closed proper concave-convex functions on $R^p \times R^q$ having domain $A^{-1}(\text{dom } K)$. Moreover,

$$\begin{aligned} \underline{H} &= \underline{KA}, & \bar{H} &= \bar{K}A, \\ \text{ri}(\text{dom } H) &= A^{-1} \text{ri}(\text{dom } K), \\ \text{cl}(\text{dom } H) &= A^{-1} \text{cl}(\text{dom } K). \end{aligned}$$

PROOF. Lemma 1.1 implies \underline{KA} and \bar{KA} are proper concave-convex functions on $R^p \times R^q$ with domain $A^{-1}(\text{dom } K)$. From Theorem 0.1(b) it is clear that a closed proper saddle-function is the least member of its equivalence class if and only if it is convex-closed. Now it follows routinely, using (6.7), (34.3) and (9.5), that \underline{KA} satisfies the six conditions of (34.3) and moreover is convex-closed. Hence \underline{KA} is closed and is the least member of its equivalence class. Similarly, \bar{KA} is closed and is the greatest member of its equivalence class. According to (34.4), two closed proper saddle-functions are equivalent if and only if they have the same kernel. Suppose $(u, v) \in \text{ri}(A^{-1}(\text{dom } K))$. By (6.7) this means $A(u, v) \in \text{ri}(\text{dom } K)$. Since \underline{K} and \bar{K} are

equivalent closed proper, $\underline{K}A(u, v) = \bar{K}A(u, v)$. This shows $\underline{K}A \sim \bar{K}A$ and hence $[\underline{K}A] = [\bar{K}A] = [H]$. If $K \in [K]$, then $\underline{K} \leq K \leq \bar{K}$ implies $\underline{K}A \leq KA \leq \bar{K}A$ and hence $KA \in [H]$ (Theorem 0.1(b)). The formulas for $ri(\text{dom } H)$ and $cl(\text{dom } H)$ are immediate from (6.7).

The next result concerns the subdifferential mapping of KA , denoted by $\partial(KA)$. The subdifferential mapping is a generalization of the gradient mapping to saddle functions which are not necessarily differentiable. For background on the subdifferential of a saddle function, we refer the reader to [48] or [43].

THEOREM 1.3. Let K and A be as in Theorem 1.2. Then

$$\partial(KA)(u, v) = A^* \partial K(A(u, v))$$

for each $(u, v) \in R^p \times R^q$, and

$$ri(A^{-1}(\text{dom } K)) \subset \text{dom } \partial(KA) \subset A^{-1}(\text{dom } K).$$

PROOF. The inclusions are immediate from (37.4) and Theorem 1.2. It follows that the identity holds trivially when $(u, v) \notin A^{-1}(\text{dom } K)$. Suppose $(u, v) \in A^{-1}(\text{dom } K)$. By the definitions, $(u^*, v^*) \in \partial(KA)(u, v)$ if and only if $u^* \in \partial(K(\cdot, A_2 v)A_1)(u)$ and $v^* \in \partial(K(A_1 u, \cdot)A_2)(v)$. Now by (34.3), $A_1 u \in \text{dom}_1 K$ implies that $K(A_1 u, \cdot)$ is a proper convex function with $ri(\text{dom } K(A_1 u, \cdot)) = ri(\text{dom}_2 K)$. Hence $\text{range } A_2 \cap ri(\text{dom}_2 K) \neq \emptyset$ and (23.9) imply that $v^* \in \partial(K(A_1 u, \cdot)A_2)(v)$ if and only if $v^* \in A_2^* \partial K(A_1 u, \cdot)(A_2 v)$, i.e. if and only if $v^* \in A_2^* \partial_2 K(A(u, v))$. Similarly, $u^* \in \partial(K(\cdot, A_2 v)A_1)(u)$ if and only if $u^* \in A_1^* \partial_1 K(A(u, v))$. The identity follows.

THEOREM 1.4. Let K and A be as in Theorem 1.2. Let f (resp. g) denote the convex (resp. concave) parent of K , and let h (resp. k) denote

the convex (resp. concave) parent of KA . Then

$$h(u, v^*) = (A_2^* f(A_1 u, \cdot))(v^*)$$

and

$$k(u^*, v) = (A_1^* g(\cdot, A_2 v))(u^*) .$$

PROOF. Suppose $u \notin \text{dom}_1 KA$. Then $h(u, \cdot)$ is constantly $+\infty$. Also, $A_1 u \notin \text{dom}_1 K$ implies $f(A_1 u, \cdot)$ is constantly $+\infty$ and hence

$$(A_2^* f(A_1 u, \cdot))(v^*) = \inf \{ f(A_1 u, y^*) \mid A_2^* y^* = v^* \} = +\infty$$

for every v^* . Now suppose $u \in \text{dom}_1 KA$. By (34.3), $A_1 u \in \text{dom}_1 K$ implies $K(A_1 u, \cdot)$ is a proper convex function with $\text{ri}(\text{dom } K(A_1 u, \cdot)) = \text{ri}(\text{dom}_2 K)$. Thus from (16.3), Theorem 0.1(a) and $\text{range } A_2 \cap \text{ri}(\text{dom}_2 K) \neq \emptyset$ it follows that $h(u, v^*) = (K(A_1 u, \cdot) A_2^*)(v^*) = (A_2^* K(A_1 u, \cdot))(v^*) = (A_2^* f(A_1 u, \cdot))(v^*)$ for every v^* .

This proves the first identity. The second is proved similarly.

COROLLARY 1.4.1. Let K and A be as in Theorem 1.2. If K is poly-
hedral, then KA is polyhedral.

PROOF. Let I_p and I_n denote the identity transformations on R^p and R^n , respectively, and let f and h be as in Theorem 1.4. Then

$h = (I_p \times A_2^*)(f(A_1 \times I_n))$. Hence (19.3.1) implies that h is polyhedral if f is polyhedral. Since KA is closed by Theorem 1.2, this concludes the proof.

The next three results concern $\text{dom}(KA)^*$. By the remark preceding Lemma 0.4, such information is of interest since it pertains to the existence of a saddle value or saddle point of KA .

COROLLARY 1.4.2. Let K and A be as in Theorem 1.2. Then

$$\text{dom}(KA)^* \subset A^* \text{dom} K^* .$$

In particular, if f (resp. g) denotes the convex (resp. concave) parent of K , then

$$\text{ri}(\text{dom}_1(KA)^*) = A_1^* \cup \{\text{ri}(\text{dom } g(\cdot, y)) \mid y \in \text{range } A_2 \cap \text{ri}(\text{dom}_2 K)\}$$

and

$$\text{ri}(\text{dom}_2(KA)^*) = A_2^* \cup \{\text{ri}(\text{dom } f(x, \cdot)) \mid x \in \text{range } A_1 \cap \text{ri}(\text{dom}_1 K)\} ,$$

where these formulas also hold when "ri" is deleted throughout.

PROOF. By Lemma 0.4.

It would be very nice for stating duality results if in addition to the inclusion of Corollary 1.4.2 we also had the inclusion

$$\text{ri}(A^* \text{dom} K^*) \subset \text{dom}(KA)^* .$$

However, this inclusion is not true in general. More discussion of this point (phrased in terms of the addition operation rather than KA) follows Corollary 3.6.3. By using recession functions, though, we are able to characterize when the above inclusion is valid. This is done in the next lemma (cf. (6.3.1)).

LEMMA 1.5. Let K and A be as in Theorem 1.2. Then for $j = 1$ and 2 ,

$$\text{cl}(\text{dom}_j(KA)^*) = \text{cl}(A_j^* \text{dom}_j K^*)$$

if and only if

$$\text{rec}_j(KA) = (\text{rec}_j K)A_j .$$

PROOF. By Theorem 0.3, $\text{rec}_2(KA) = \delta^*(\cdot \mid \text{dom}_2(KA)^*)$ and $(\text{rec}_2 K)A_2 = \delta^*(A_2 \cdot \mid \text{dom}_2 K^*)$. Now apply (16.3.1) and (13.1.1). The assertion for $j = 1$ is proved similarly.

LEMMA 1.6. Let K and A be as in Theorem 1.2. Then

$$(\text{rec}_1 KA)(u) = \inf\{(\text{rec } K(\cdot, y))(A_1 u) \mid y \in \text{range } A_2 \cap \text{ri}(\text{dom}_2 K)\}$$

and

$$(\text{rec}_2 KA)(v) = \sup\{(\text{rec } K(x, \cdot))(A_2 v) \mid x \in \text{range } A_1 \cap \text{ri}(\text{dom}_1 K)\} .$$

PROOF. By definition and Theorem 1.2,

$$\begin{aligned} (\text{rec}_2 KA)(v) &= \sup\{(\text{rec}(KA)(u, \cdot))(v) \mid u \in A_1^{-1} \text{ri}(\text{dom}_1 K)\} . \\ &= \sup\{(\text{rec } K(x, \cdot)A_2)(v) \mid x \in \text{range } A_1 \cap \text{ri}(\text{dom}_1 K)\} . \end{aligned}$$

If $x \in \text{ri}(\text{dom}_1 K)$, then (34.3) and (9.5) imply $(\text{rec } K(x, \cdot)A_2)(v) = (\text{rec } K(x, \cdot))(A_2 v)$.

This proves one formula, and the other is proved similarly.

Now we develop another operation promised, an operation which is analogous to that of taking the image Af of a convex function f under a linear transformation A . Suppose K is a concave-convex function on $R^m \times R^n$ and $A = A_1 \times A_2$ is a linear transformation from $R^m \times R^n$ to $R^p \times R^q$. We seek a condition on K and A ensuring that all the functions on $R^p \times R^q$ either of the form

$$(u, v) \rightarrow \sup_{\{x \mid A_1 x = u\}} \inf_{\{y \mid A_2 y = v\}} \tilde{K}(x, y)$$

or of the form

$$(u, v) \rightarrow \inf_{\{y \mid A_2 y = v\}} \sup_{\{x \mid A_1 x = u\}} \tilde{K}(x, y) ,$$

for $\tilde{K} \in [K]$, belong to a single equivalence class of concave-convex functions on $R^p \times R^q$. (By (5.5) and (5.7), these functions are indeed concave-convex.)

By analogy with the operation in the convex function case, this equivalence class (when it exists) will usually be denoted by $[AK]$. Theorem 1.8 gives a condition which guarantees that $[AK]$ exists and, moreover, that all of its members are closed and proper, and that its conjugate is $[K^* A^*]$.

The proof of Theorem 1.8 relies on the following technical result. For simplicity we "dualize" the notation used up until this point in §1, i.e., we write K in place of K^* and A in place of A^* .

LEMMA 1.7. Let K be a closed proper concave-convex function on $R^m \times R^n$, and let $A = A_1 \times A_2$ be a linear transformation from $R^m \times R^n$ to $R^p \times R^q$. Assume $\text{range } A^* \cap \text{ri}(\text{dom } K^*) \neq \emptyset$. Then

$$(\overline{K^* A^*})^*(u, v) = \sup\{\text{cl}(A_2 K(x, \cdot))(v) \mid A_1 x = u\},$$

where the supremum can be taken over just those x in $\text{dom}_1 K$ such that $A_1 x = u$, and

$$(\underline{K^* A^*})^*(u, v) = \inf\{\text{cl}(A_1 \overline{K}(\cdot, y))(u) \mid A_2 y = v\},$$

where the infimum can be taken over just those y in $\text{dom}_2 K$ such that $A_2 y = v$.

PROOF. Only the first formula will be proved, as the second can be proved similarly. Let J denote the lower conjugate of $\overline{K^* A^*}$. The definitions yield

$$J(u, v) = \sup_{v^*} \{ \langle v^*, v \rangle + \inf_{u^*} \{ \langle u^*, u \rangle - (\overline{K^*}(\cdot, A_2^* v^*) A_1^*)(u^*) \} \}.$$

Since $\overline{K^*}$ is concave-closed, it follows from (34.3) and (6.3.1) that $\text{ri}(\text{dom } \overline{K^*}(\cdot, y^*))$ equals $\text{ri}(\text{dom}_1 K^*)$ when $y^* \in \text{dom}_2 K^*$ and equals R^m when $y^* \notin \text{dom}_2 K^*$. Hence (16.3) and the hypothesis $\text{range } A_1^* \cap \text{ri}(\text{dom}_1 K^*) \neq \emptyset$ imply

$(\bar{K}^*(\cdot, A_2^{**} A_1^*)^*)(u) = (A_1 \bar{K}^*(\cdot, A_2^{**})^*)(u) = \sup\{k(x, A_2^{**}) \mid A_1 x = u\}$ for every v^* , where k denotes the concave parent of K^* . Thus,

$$\begin{aligned} J(u, v) &= \sup_{v^*} \{ \langle v^*, v \rangle + \sup_{x \in A_1^{-1} u} k(x, A_2^{**}) \} \\ &= \sup_{x \in A_1^{-1} u} \sup_{v^*} \{ \langle v^*, v \rangle - (-k)(x, A_2^{**}) \}. \end{aligned}$$

But Theorem 0.2 implies $-k$ is the convex parent f of K , and hence (16.3) implies $\sup_{v^*} \{ \langle v^*, v \rangle - (-k)(x, A_2^{**}) \} = (f(x, \cdot) A_2^*)^*(v) = \text{cl}(A_2 f(x, \cdot))^*(v) = \text{cl}(A_2 \underline{K}(x, \cdot))(v)$. This establishes the asserted formula for J . Finally, for each $x \notin \text{dom}_1 K$, the fact that \underline{K} is convex-closed implies $\underline{K}(x, \cdot)$ and hence $\text{cl}(A_2 \underline{K}(x, \cdot))$ is constantly $-\infty$.

THEOREM 1.8. Let K and A be as in Lemma 1.7 and assume
 $\text{range } A^* \cap \text{ri}(\text{dom } K^*) \neq \emptyset$. Define functions J_1 and J_2 on $R^p \times R^q$ by

$$J_1(u, v) = \sup_{\{x \mid A_1 x = u\}} \inf_{\{y \mid A_2 y = v\}} K(x, y)$$

and

$$J_2(u, v) = \inf_{\{y \mid A_2 y = v\}} \sup_{\{x \mid A_1 x = u\}} \bar{K}(x, y).$$

Then there exists an equivalence class $[AK]$ which contains every concave-convex function J on $R^p \times R^q$ satisfying $J_1 \leq J \leq J_2$. Moreover, $[AK]$ is closed and proper and its conjugate is $[K^* A^*]$. If $[K]$ is polyhedral, then $[AK]$ is polyhedral.

PROOF. Theorem 1.2 implies that $\underline{K}^* A^*$ and $\overline{K}^* A^*$ belong to a closed proper equivalence class $[K^* A^*]$. Let $[AK]$ denote the conjugate equivalence class. From Lemma 1.7 and the fact that $cl f \leq f$ for any convex function f and $cl g \geq g$ for any concave function g , it follows that

$$(\overline{K}^* A^*)^* \leq J_1 \text{ and } J_2 \leq (\underline{K}^* A^*)^* .$$

Hence Theorem 0.1(b) implies that each concave-convex function J on $R^p \times R^q$ satisfying $J_1 \leq J \leq J_2$ belongs to $[AK]$. The polyhedral assertion follows from Corollary 1.4.1 and the fact that K^* is polyhedral whenever K is.

Recall from (37.5) that ∂K^* is just the inverse of ∂K , considered as a multivalued mapping. By combining this fact with Theorem 1.3 and the fact that $[AK]$ is the conjugate of $[K^* A^*]$ (Theorem 1.8), we immediately have a characterization of the subdifferential $\partial(AK)$ of AK .

The basic hypothesis used throughout this section is dualized in the following lemma.

LEMMA 1.9. Let K and A be as in Lemma 1.7. Then for $j = 1$ and 2 the following conditions are equivalent:

- (a_j) $\text{range } A_j^* \cap \text{ri}(\text{dom}_j K^*) \neq \emptyset$;
- (b_j) $A_j^{-1}\{0\} \cap (\text{rec cone}_j K)$ is a subspace ;
- (c_j) $A_j^{-1}\{0\} \cap (\text{rec cone}_j K) \subset -(\text{rec cone}_j K)$.

PROOF. Apply Lemma 0.7 with $L_j = \text{range } A_j^*$.

We conclude this section with two examples showing that Theorems 1.2 and 1.8 can fail if their hypotheses are weakened. These examples are presented in the notational scheme of Theorem 1.2.

EXAMPLE 1.10. Take $m = n = p = q = 1$, and let A_1 and A_2 each be the zero transformation on R . Let K be a member of the equivalence class of closed proper concave-convex functions on $R \times R$ having as kernel the function

$$(u, v) \rightarrow u^v, \quad \forall u, v \in (0, 1) \times (0, 1).$$

(This equivalence class is discussed in [48, p. 360].) Then $\text{dom } K = [0, 1] \times [0, 1]$, $\underline{K}(0, 0) = 0$, $\bar{K}(0, 0) = 1$, and $\underline{K}(u, v) = u^v = \bar{K}(u, v)$ whenever $(u, v) \in \text{dom } K \setminus \{(0, 0)\}$. Moreover, for each $\alpha \in [0, 1]$ the function K_α belongs to $[K]$, where $K_\alpha(0, 0) = \alpha$ and $K_\alpha(u, v) = \underline{K}(u, v)$ whenever $(u, v) \neq (0, 0)$. Observe that, for $j = 1$ and 2 , $\text{range } A_j \cap \text{dom}_j K \neq \emptyset$ while $\text{range } A_j \cap \text{ri}(\text{dom}_j K) = \emptyset$. Also, for any $\tilde{K} \in [K]$, the function $\tilde{K}A$ is constantly equal to $\tilde{K}(0, 0)$. Since $0 \leq \tilde{K}(0, 0) \leq 1$, this implies that $\tilde{K}A$ is closed and proper. However, it also implies that, for any two elements K_1 and K_2 of $[K]$, K_1A is equivalent to K_2A if and only if $K_1(0, 0) = K_2(0, 0)$. Thus, as \tilde{K} ranges over $[K]$ the functions $\tilde{K}A$ determine 2^{R_0} distinct equivalence classes of closed proper saddle-functions (cf. Theorem 1.2). Now let J_1 and J_2 be defined as in Theorem 1.8 (except for dualizing the notation). Since A_j^* is the zero transformation on R ,

$$J_1(u^*, v^*) = \begin{cases} \sup_R \inf_R \underline{K}^* & \text{if } u^* = 0 \text{ and } v^* = 0 \\ +\infty & \text{if } u^* = 0 \text{ and } v^* \neq 0 \\ -\infty & \text{if } u^* \neq 0 \end{cases}$$

and

$$I_2(u^\circ, v^\circ) = \begin{cases} \inf_R \sup_R \bar{K}^\circ & \text{if } u^\circ = 0 \text{ and } v^\circ = 0 \\ -\infty & \text{if } u^\circ \neq 0 \text{ and } v^\circ = 0 \\ +\infty & \text{if } v^\circ \neq 0. \end{cases}$$

But $\sup_R \inf_R \underline{K}^\circ = -(\underline{K}^\circ)^\circ(0, 0) = -\bar{K}(0, 0) = -1$, and similarly $\inf_R \sup_R \bar{K}^\circ = 0$.

Hence I_1 and I_2 are closed and proper but not equivalent (cf. Theorem 1.8).

EXAMPLE 1.11. Let K and A_1 be as in Example 1.10, but now let A_2 be the identity transformation on R . Observe that $\text{range } A_2 \cap \text{ri}(\text{dom}_2 K) \neq \emptyset$ and $\text{range } A_1 \cap \text{dom}_1 K \neq \emptyset$ but $\text{range } A_1 \cap \text{ri}(\text{dom}_1 K) = \emptyset$. For each $\bar{K} \in [K]$,

$$\bar{K}A(u, v) = \bar{K}(0, v) = \begin{cases} 0 & \text{if } v \in (0, 1) \\ \bar{K}(0, 0) & \text{if } v = 0 \\ +\infty & \text{if } v \notin [0, 1], \end{cases}$$

where $0 \leq \bar{K}(0, 0) \leq 1$. This implies that $\bar{K}A$ is proper with domain $R \times [0, 1]$ and that $\bar{K}A$ is closed if and only if $\bar{K}(0, 0) = 0$. It also implies that, for any two elements K_1 and K_2 of $[K]$, K_1A is equivalent to K_2A if and only if $K_1(0, 0) = K_2(0, 0)$. Recalling from Example 1.10 the functions K_α for $\alpha \in [0, 1]$, we conclude that as \bar{K} ranges over $[K]$ the functions $\bar{K}A$ determine 2^{\aleph_0} distinct equivalence classes of proper saddle functions, where only the class containing $\underline{K}A$ is closed (cf. Theorem 1.2).

§2. Sharper Results

In this section the results of §1 concerning $[AK]$ are sharpened. In particular, conditions are given under which (i) the effective domain of AK cannot "collapse" significantly from $A \text{ dom } K$; (ii) the extrema appearing in the definitions of the elements J_1 and J_2 of $[AK]$ are actually attained by saddle points; and (iii) the saddle functions J_1 and J_2 actually coincide with the least and greatest elements of $[AK]$. This is done in Theorems 2.4 and 2.5. Lemma 2.6 states some simple conditions sufficient for the more general hypotheses of Theorems 2.4 and 2.5 to hold.

Throughout §2 we adopt the notational setting of Theorem 1.8. That is, K is a closed proper concave-convex function on $R^m \times R^n$, $A = A_1 \times A_2$ is a linear transformation from $R^m \times R^n$ to $R^p \times R^q$, and J_1 and J_2 are functions defined on $R^p \times R^q$ by

$$J_1(u, v) = \sup_{\{x | A_1 x = u\}} \inf_{\{y | A_2 y = v\}} K(x, y)$$

and

$$J_2(u, v) = \inf_{\{y | A_2 y = v\}} \sup_{\{x | A_1 x = u\}} \bar{K}(x, y).$$

Theorem 2.4 rests on Theorem 2.2, which in turn relies on the following technical lemma.

LEMMA 2.1. Let f be a proper convex function on R^n , let D be a convex set such that $D \subset \text{dom } f \subset \text{cl } D$, and let E be a convex set such that $E \cap \text{ri } D \neq \emptyset$. Then

$$\inf_{E \cap D} f = \inf_E f.$$

PROOF. By (6.3.1), $D \subset \text{dom } f \subset \text{cl } D$ implies $\text{ri } D = \text{ri}(\text{dom } f)$. Hence $S = E \cap \text{ri}(\text{dom } f) \subset E \cap D \subset E$ implies trivially that

$$\inf_S f \geq \inf_{E \cap D} f \geq \inf_E f.$$

Let $y \in E$ be given. If $y \notin \text{dom } f$, then $f(y) = +\infty \geq \inf_S f$. Suppose $y \in \text{dom } f$. Since $E \cap \text{ri } D \neq \emptyset$, we can pick an $x \in S$. Then (6.1) implies that $y_\lambda = (1-\lambda)x + \lambda y \in S$ for each $0 \leq \lambda < 1$. Hence (7.5) implies that $f(y) \geq \text{cl } f(y) = \lim_{\lambda \uparrow 1} f(y_\lambda) \geq \inf\{f(y_\lambda) \mid 0 \leq \lambda < 1\} \geq \inf_S f$. This shows that $f(y) \geq \inf_S f$ for every $y \in E$. Thus $\inf_E f \geq \inf_S f$, and the proof is complete.

THEOREM 2.2. Let $(u, v) \in A \text{ ri}(\text{dom } K)$ and assume that

$$A_2^{-1}(0) \cap \bigcap \{ \text{rec cone } \underline{K}(x, \cdot) \mid x \in \text{ri}(\text{dom}_1 K), A_1 x = u \}$$

and

$$A_1^{-1}(0) \cap \bigcap \{ \text{rec cone } \bar{K}(\cdot, y) \mid y \in \text{ri}(\text{dom}_2 K), A_2 y = v \}$$

are subspaces. Then there exists a nonempty closed convex product set $X \times Y$ in $\text{dom } K \cap A^{-1}((u, v))$ such that $(x, y) \in X \times Y$ if and only if (x, y) is a saddle point of K with respect to $A^{-1}((u, v))$ for each $\bar{K} \in [K]$. If the two sets in the hypothesis are actually nullspaces, then $X \times Y$ is bounded.

PROOF. Define a concave-convex function L on $\mathbb{R}^m \times \mathbb{R}^n$ by

$$L(x, y) = \begin{cases} 0 & \text{if } A_1 x = u \text{ and } A_2 y = v \\ +\infty & \text{if } A_1 x = u \text{ and } A_2 y \neq v \\ -\infty & \text{if } A_1 x \neq u. \end{cases}$$

Clearly, L is closed and its effective domain is $A^{-1}((u, v))$. Since $(u, v) \in A \operatorname{ri}(\operatorname{dom} K)$, $\operatorname{ri}(\operatorname{dom} K) \cap \operatorname{ri}(\operatorname{dom} L) \neq \emptyset$. Therefore Theorem 3.2 (which doesn't depend on the results of §2) implies that the equivalence class $[K] + [L]$ is defined and has domain $S \times T$, where $S = A_1^{-1}(u) \cap \operatorname{dom}_1 K$ and $T = A_2^{-1}(v) \cap \operatorname{dom}_2 K$. Moreover, Theorem 3.2 also implies that for any $\tilde{K} \in [K]$, $[K] + [L]$ contains the closed proper saddle function M given by

$$M(x, y) = \begin{cases} \tilde{K}(x, y) & \text{if } x \in S \text{ and } y \in T \\ +\infty & \text{if } x \in S \text{ and } y \notin T \\ -\infty & \text{if } x \notin S \end{cases}$$

Suppose $x \in \operatorname{ri} S = A_1^{-1}(u) \cap \operatorname{ri}(\operatorname{dom}_1 K)$ (use (6.5)). Then (34.3) implies $\tilde{K}(x, \cdot) = \underline{K}(x, \cdot)$ is a closed proper convex function with effective domain $\operatorname{dom}_2 K$. Hence (9.3) and the definition of M imply

$$\operatorname{rec} M(x, \cdot) = \operatorname{rec} \underline{K}(x, \cdot) + \operatorname{rec} \delta(\cdot | A_2^{-1}(v)) .$$

But $\operatorname{rec} \delta(\cdot | A_2^{-1}(v)) = \delta(\cdot | A_2^{-1}(0))$. Therefore

$$\operatorname{rec} M(x, \cdot) = A_2^{-1}(0) \cap \operatorname{rec} \operatorname{cone} \underline{K}(x, \cdot) .$$

Since $M(x, \cdot) = \underline{M}(x, \cdot)$ whenever $x \in \operatorname{ri} S$ (Theorem 0.1(b)), this implies that

$$\operatorname{rec} \operatorname{cone}_2 M = A_2^{-1}(0) \cap \bigcap \{ \operatorname{rec} \operatorname{cone} \underline{K}(x, \cdot) \mid x \in \operatorname{ri} S \} .$$

By hypothesis this is a subspace. Similarly, $\operatorname{rec} \operatorname{cone}_1 M$ is a subspace. It follows from Lemma 0.5 that $(0, 0) \in \operatorname{ri}(\operatorname{dom} M^\circ)$, and hence (37.5.3) implies that $\partial M^\circ(0, 0)$ is a nonempty closed convex product set $X \times Y$. By (37.5), $\partial M^\circ(0, 0) \subset \operatorname{dom} \partial M$. But Theorem 3.9 (which doesn't depend on the results of

(2) implies that $\text{dom } \partial M = \text{dom } \partial K \cap \text{dom } \partial L$, and (37.4) implies $\text{dom } \partial L \subset \text{dom } L$. Therefore $X \times Y \subset \text{dom } \partial K \cap A^{-1}((u, v))$. Now $(x, y) \in X \times Y$ if and only if (x, y) is a saddle point of M , which (by (36.3)) occurs if and only if (x, y) is a saddle point of \tilde{K} with respect to $S \times T$. Using $(x, y) \in \text{dom } K$ together with (34.3) and Lemma 2.1, it follows that this occurs if and only if (x, y) is a saddle point of \tilde{K} with respect to $A^{-1}((u, v))$. Since any member of $[K]$ could be taken in the definition of M , (36.4) implies that $(x, y) \in X \times Y$ if and only if (x, y) is a saddle point of \tilde{K} with respect to $A^{-1}((u, v))$ for each $\tilde{K} \in [K]$. Finally, suppose the sets in the hypothesis are actually nullspaces. Then $\text{rec cone}_j M = \{0\}$ for $j = 1$ and 2 , so that Lemma 0.5 implies $(0, 0) \in \text{int}(\text{dom } M^\circ)$. From this, (34.3) and (23.4) it follows that the sets $\partial M^\circ(\cdot, 0)(0) = X$ and $\partial M^\circ(0, \cdot)(0) = Y$ are bounded.

The conditions used in Theorems 2.4 and 2.5 are given in the next lemma. In contrast to the "global" conditions used in Theorem 1.8, these are pointwise in character.

LEMMA 2.3. For $x \in \text{dom}_1 K$ the following three conditions are equivalent, and they imply $A_1 x \in \text{dom}_1 J_1$:

- (a₁) $\text{range } A_2^\circ \cap \text{ri}(\text{dom } K(x, \cdot))^\circ \neq \emptyset$;
- (a₂) $A_2^{-1}(0) \cap \text{rec cone } \underline{K}(x, \cdot)$ is a subspace;
- (a₃) $A_2^{-1}(0) \cap \text{rec cone } \underline{K}(x, \cdot) \subset -\text{rec cone } \underline{K}(x, \cdot)$.

Similarly, for $y \in \text{dom}_2 K$ the following three conditions are equivalent and they imply $A_2 y \in \text{dom}_2 J_2$:

$$(b_1) \quad \text{range } A_1^\circ \cap \text{ri}(\text{dom } K(\cdot, y)^\circ) \neq \emptyset;$$

$$(b_2) \quad A_1^{-1}\{0\} \cap \text{rec cone } \bar{K}(\cdot, y) \text{ is a subspace};$$

$$(b_3) \quad A_1^{-1}\{0\} \cap \text{rec cone } \bar{K}(\cdot, y) \subset -\text{rec cone } \bar{K}(\cdot, y).$$

PROOF. Only the first assertion is proved, as the second can be proved similarly. Since $\text{rec cone } \underline{K}(x, \cdot)$ is a convex cone, clearly (a_2) holds if and only if (a_3) holds. By Theorem 0.1, $K(x, \cdot)^\circ$ is proper convex and its conjugate is $\underline{K}(x, \cdot)$. Hence (16.2.1) implies that (a_1) fails if and only if (a_3) fails. Thus, the three conditions $(a_1) - (a_3)$ are equivalent. Suppose now that x satisfies (a_3) . Since $\underline{K}(x, \cdot)$ is closed proper convex, (9.2) implies that $A_2 \underline{K}(x, \cdot)$ is too. Hence $A_2 \underline{K}(x, \cdot)$ is never $-\infty$. But $A_2 \underline{K}(x, \cdot) \leq J_1(A_1 x, \cdot)$. Therefore $A_1 x \in \text{dom } J_1$.

THEOREM 2.4. Assume that each $x \in \text{ri}(\text{dom}_1 K)$ (resp. $y \in \text{ri}(\text{dom}_2 K)$) satisfies one of the equivalent conditions (a_1) (resp. (b_1)) of Lemma 2.3. Then the conclusions of Theorem 1.8 hold, and

$$\text{ri}(\text{Adom } K) \subset \text{dom } AK \subset \text{Adom } K.$$

Furthermore, for each $(u, v) \in \text{ri}(\text{Adom } K)$ there exists a nonempty closed convex product set $X \times Y$ in $\text{dom } \partial K \cap A^{-1}\{(u, v)\}$ such that

(1) $(x, y) \in X \times Y$ if and only if (x, y) is a saddle point of \tilde{K} with respect to $A^{-1}\{(u, v)\}$ for each $\tilde{K} \in [K]$, and

(2) $(x, y) \in X \times Y$ implies $\tilde{J}(u, v) = \tilde{K}(x, y)$ for every $\tilde{J} \in [AK]$ and $\tilde{K} \in [K]$.

PROOF. The hypothesis implies that $A_1^{-1}\{0\} \cap \text{rec cone}_1 K$ is a subspace

for $j = 1$ and 2 . Hence by Lemma 1.9 the conclusions of Theorem 1.8 hold, and in particular J_1 and J_2 belong to $[AK]$, where $[(AK)^*] = [K^*A^*]$. Thus $\text{dom } AK = \text{dom}_1 J_1 \times \text{dom}_2 J_2$. Therefore the hypothesis and Lemma 2.3 imply that $\text{ri}(A \text{ dom } K) \subset \text{dom } AK$. On the other hand, Lemma 1.9 and Corollary 1.4.2 imply that $\text{dom } AK \subset A \text{ dom } K$. Let $(u, v) \in \text{ri}(A \text{ dom } K)$. By Theorem 2.2 there exists a nonempty closed convex product set $X \times Y$ in $\text{dom } \partial K \cap A^{-1}\{(u, v)\}$ such that (1) holds. Suppose $(x, y) \in X \times Y$. Since (1) implies (x, y) is a saddle point of \underline{K} with respect to $A^{-1}\{(u, v)\}$, certainly $J_1(u, v) = \underline{K}(x, y)$. Since $\text{ri}(A \text{ dom } K) = \text{ri}(\text{dom } AK)$ by (6.3.1), Theorem 0.1(b) implies that $J_1(u, v) = \tilde{J}(u, v)$ for each $\tilde{J} \in [AK]$. Also, $(x, y) \in \text{dom } \partial K$ and (37.4.1) imply that $\underline{K}(x, y) = \tilde{K}(x, y)$ for each $\tilde{K} \in [K]$. This establishes (2).

THEOREM 2.5. Assume that each $x \in \text{dom}_1 K$ (resp. $y \in \text{dom}_2 K$) satisfies one of the equivalent conditions (a₁) (resp. (b₁)) of Lemma 2.3. Then $\text{dom } AK$ actually equals $A \text{ dom } K$. Moreover, writing $\text{cl}_2(AK) = \underline{J}$ and $\text{cl}_1(AK) = \bar{J}$,

$$J_1 = \underline{J} \text{ and } J_2 = \bar{J} \text{ on range } A.$$

In particular, $J_1(u, v) = \underline{J}(u, v)$ except when $u \in \text{range } A_1 \setminus A_1 \text{ dom}_1 K$ and $v \notin \text{range } A_2$, and $J_2(u, v) = \bar{J}(u, v)$ except when $u \notin \text{range } A_1$ and $v \in \text{range } A_2 \setminus A_2 \text{ dom}_2 K$.

PROOF. By Lemma 2.3, $A \text{ dom } K \subset \text{dom}_1 J_1 \times \text{dom}_2 J_2$. From this it follows as in the proof of Theorem 2.4 that $A \text{ dom } K = \text{dom } AK$. We only prove the assertion about J_1 , as the other is similar. From the proof of Lemma 2.3, $A_2 \underline{K}(x, \cdot)$ is closed for each $x \in \text{dom}_1 K$. Hence Lemma 1.7 implies

$$\underline{J}(u, v) = \sup \{A_2 \underline{K}(x, \cdot)(v) \mid x \in \text{dom}_1 K, A_1 x = u\}. \quad (1)$$

If $x \notin \text{dom}_1 K$, then $\underline{K}(x, \cdot)$ is constantly $-\infty$, so that

$$\underline{A}_2 \underline{K}(x, \cdot)(v) = \begin{cases} -\infty & \text{if } v \in \text{range } \underline{A}_2 \\ +\infty & \text{if } v \notin \text{range } \underline{A}_2 \end{cases} \quad (2)$$

Since $J_1(u, v) = \sup\{\underline{A}_2 \underline{K}(x, \cdot)(v) \mid \underline{A}_1 x = u\}$ by definition, equation (1) implies that $J_1(u, v)$ equals

$$\max\{J(u, v), \sup\{\underline{A}_2 \underline{K}(x, \cdot)(v) \mid \underline{A}_1 x = u, x \notin \text{dom}_1 K\}\},$$

which by equation (2) equals $J(u, v)$ whenever $v \in \text{range } \underline{A}_2$. Henceforth assume $v \notin \text{range } \underline{A}_2$. Suppose $u \in \underline{A}_1 \text{dom}_1 K$. Pick an $x \in \text{dom}_1 K$ such that $\underline{A}_1 x = u$. Since $\underline{A}_2^{-1}\{v\} = \emptyset$, $+\infty = \underline{A}_2 \underline{K}(x, \cdot)(v) \leq J(u, v) \leq J_1(u, v)$. Hence $J(u, v) = J_1(u, v) = +\infty$ whenever $u \in \underline{A}_1 \text{dom}_1 K$. Observe also that the convention $\sup \emptyset = -\infty$ implies $J(u, v) \leq J_1(u, v) = -\infty$ whenever $u \notin \text{range } \underline{A}_1$. In the only remaining case, i. e. when $u \in \text{range } \underline{A}_1 \setminus \underline{A}_1 \text{dom}_1 K$, equation (1) implies $J(u, v) = \sup \emptyset = -\infty$ while $J_1(u, v) = \sup\{\inf \emptyset \mid \underline{A}_1 x = u\} = +\infty$.

While the hypotheses of Theorems 2.4 and 2.5 are general, they may appear somewhat cumbersome to check. The next lemma gives simpler, "global" conditions on K and A which imply the hypotheses of both Theorems 2.4 and 2.5. Note that these conditions are met, for example, when $\text{dom } K$ is bounded.

For a nonempty convex set C in \mathbb{R}^n , define the recession cone of C to be the set

$$0^+ C = \{y \mid x + \lambda y \in C, \forall x \in C \forall \lambda \geq 0\}.$$

LEMMA 2.6. The three following conditions are equivalent, and they imply that conditions (a₁) - (a₃) of Lemma 2.3 hold for each $x \in \text{dom}_1 K$:

$$(c_1) \quad A_2^{-1}\{0\} \cap 0^+ \text{cl}(\text{dom}_2 K) = \{0\};$$

$$(c_2) \quad A_2^{-1}\{v\} \cap \text{dom}_2 K \text{ is bounded for each } v \in R^q;$$

$$(c_3) \quad A_2^{-1}\{v\} \cap \text{ri}(\text{dom}_2 K) \text{ is nonempty and bounded for some } v \in R^q.$$

Similarly, the three following conditions are equivalent, and they imply that conditions $(b_1) - (b_3)$ of Lemma 2.3 hold for each $y \in \text{dom}_2 K$:

$$(d_1) \quad A_1^{-1}\{0\} \cap 0^+ \text{cl}(\text{dom}_1 K) = \{0\};$$

$$(d_2) \quad A_1^{-1}\{u\} \cap \text{dom}_1 K \text{ is bounded for each } u \in R^p;$$

$$(d_3) \quad A_1^{-1}\{u\} \cap \text{ri}(\text{dom}_1 K) \text{ is nonempty and bounded for some } u \in R^p.$$

PROOF. Only the first assertion is proved, since the second is similar.

For each $v \in A_2 \text{dom}_2 K$, (8.3.3) and (8.4) imply that

$$A_2^{-1}\{v\} \cap \text{cl}(\text{dom}_2 K) \text{ is bounded} \iff A_2^{-1}\{0\} \cap 0^+ \text{cl}(\text{dom}_2 K) = \{0\}. \quad (1)$$

It follows from this that (c_1) implies (c_2) . By picking any $v \in A_2 \text{ri}(\text{dom}_2 K)$

it follows that (c_2) implies (c_3) . Now assume (c_3) . Then (6.3.1) and

(6.5.1) imply that

$$A_2^{-1}\{v\} \cap \text{cl}(\text{dom}_2 K) = A_2^{-1}\{v\} \cap \text{cl}(\text{ri}(\text{dom}_2 K)) = \text{cl}(A_2^{-1}\{v\} \cap \text{ri}(\text{dom}_2 K)).$$

That this set is bounded follows from the fact that $A_2^{-1}\{v\} \cap \text{ri}(\text{dom}_2 K)$ is

bounded. Hence (c_1) follows by the equivalence (1) above. Finally, let

$x \in \text{dom}_1 K$ be given. Write $f = \underline{K}(x, \cdot)$ and $C = \text{dom } f$. Then by (34.3) and

(6.3.1), f is a proper convex function with $\text{cl } C = \text{cl}(\text{dom}_2 K)$. But by (8.5)

and (8.1) it follows easily that $\text{dom}(\text{rec } f) \subset 0^+ C \subset 0^+(\text{cl } C)$. Hence $\text{rec cone } \underline{K}(x, \cdot) \subset 0^+(\text{cl dom}_2 K)$, and therefore (c_1) implies that x satisfies (a_2) of Lemma 3.3.

Finally, we remark that when conditions (c_1) and (d_1) of Lemma 2.6 are met, the sets X and Y given by Theorem 2.4 for each $(u, v) \in \text{Ari}(\text{dom } K)$ are actually bounded and hence compact. This is because the two sets in the hypothesis of Theorem 2.2 are then nullspaces.

§3. Addition and Minimax Convolution

In this section the operation of addition for equivalence classes of saddle functions is developed, and along with it a new operation called minimax convolution. It is shown that these two operations are dual to each other with respect to the basic conjugacy correspondence. The results obtained closely parallel known theorems about the dual operations of addition and infimal convolution of convex functions.

Both these operations and the duality relationship between them can be developed in a manner exactly parallel to §§1 and 2. However, to avoid such essentially repetitive proofs we instead develop only the addition operation separately and then obtain the rest of the results as special cases of those of §§1 and 2. This entails defining separable saddle functions and proving some technical facts concerning them. Finally, the section concludes with an example which leads to a conjecture about maximal monotone operators.

There are two technical difficulties involved in defining the operation of addition. The first stems from the fact that we are working with extended-real-valued functions; we must deal somehow with the expression $\infty - \infty$. The second and more fundamental difficulty is that, from the point of view of minimax theory, we want to define addition of whole equivalence classes and not just individual functions. The following definition is designed to handle both these difficulties.

For $i = 1, \dots, s$ let K_i be a concave-convex function on $R^m \times R^n$ with effective domain $C_i \times D_i$. We say that $[K_1] + \dots + [K_s]$ is defined if and only if the sets $C = C_1 \cap \dots \cap C_s$ and $D = D_1 \cap \dots \cap D_s$ are nonempty and

$$\tilde{K}_1(x, y) + \dots + \tilde{K}_s(x, y) = K_1(x, y) + \dots + K_s(x, y)$$

whenever $(x, y) \in \text{ri } C \times \text{ri } D$ and $\tilde{K}_1 \in [K_1], \dots, \tilde{K}_s \in [K_s]$. In this event $[K_1] + \dots + [K_s]$, which will usually be written as $[K_1 + \dots + K_s]$, is defined to be the unique equivalence class of closed proper concave-convex functions on $R^m \times R^n$ having as kernel the function on $\text{ri } C \times \text{ri } D$ given by

$$(x, y) \rightarrow K_1(x, y) + \dots + K_s(x, y).$$

Such a unique equivalence class exists by (34.5.1). The operation which sends $[K_1], \dots, [K_s]$ into $[K_1 + \dots + K_s]$ is, quite naturally, called addition.

LEMMA 3.1. For $i = 1, \dots, s$ let K_i be a closed proper concave-convex function on $R^m \times R^n$ with effective domain $C_i \times D_i$. Then $[K_1] + \dots + [K_s]$ is defined if either $C_1 \cap \dots \cap C_s \neq \emptyset$ and $\text{ri } D_1 \cap \dots \cap \text{ri } D_s \neq \emptyset$ or $\text{ri } C_1 \cap \dots \cap \text{ri } C_s \neq \emptyset$ and $D_1 \cap \dots \cap D_s \neq \emptyset$.

PROOF. This follows easily from (6.5) and Theorem 0.1(b).

It is actually not hard to establish a weaker condition sufficient for $[K_1] + \dots + [K_s]$ to be defined. Loosely speaking, the condition is just that the K_i be closed and that (possibly after renumbering the K_i 's) there exist an integer r , $0 \leq r \leq s$, such that

$$\text{ri } C_1 \cap \dots \cap \text{ri } C_r \cap C_{r+1} \cap \dots \cap C_s \neq \emptyset$$

and

$$D_1 \cap \dots \cap D_r \cap \text{ri } D_{r+1} \cap \dots \cap \text{ri } D_s \neq \emptyset.$$

(The conditions in Lemma 3.1 correspond to the values $r = 0$ and $r = s$).

Instead of appealing to (6.5), the proof uses the generalization of (6.5) given as Lemma A.6 in the Appendix.

THEOREM 3.2. Let K_1, \dots, K_s be closed proper concave-convex functions on $R^m \times R^n$ such that $\text{ri}(\text{dom } K_1) \cap \dots \cap \text{ri}(\text{dom } K_s) \neq \emptyset$. Then $[K_1] + \dots + [K_s]$ is defined, has effective domain $\text{dom } K_1 \cap \dots \cap \text{dom } K_s$, and contains the closed proper saddle function K given by

$$K(x, y) = \begin{cases} \sum K_i(x, y) & \text{if } x \in C \text{ and } y \in D \\ +\infty & \text{if } x \in C \text{ and } y \notin D \\ -\infty & \text{if } x \notin C \end{cases}$$

PROOF. Lemma 3.1 implies $[K_1] + \dots + [K_s]$ is defined. Hence it is the unique equivalence class of closed proper concave-convex functions on $R^m \times R^n$ having the same kernel as K . Therefore by (34.4) the proof will be complete once we show K is closed. This we do by checking that K satisfies the six conditions of (34.3). This follows routinely by applying (34.3) to the K_i 's with the aid of (6.5).

In order to apply the results of §§1 and 2 to an equivalence class $[K_1 + \dots + K_s]$ and its conjugate, we need to define and establish some properties of "separable" saddle functions. For $i = 1, \dots, s$ let K_i be a proper concave-convex function on $R^{m_i} \times R^{n_i}$ with effective domain $C_i \times D_i$. Write $m = \sum m_i$, $n = \sum n_i$ and define a function (K_1, \dots, K_s) on $R^m \times R^n$ by

$$(K_1, \dots, K_s)(x, y) = \begin{cases} \sum K_i(x_i, y_i) & \text{if } x \in C \text{ and } y \in D \\ +\infty & \text{if } x \in C \text{ and } y \notin D \\ -\infty & \text{if } x \notin C \end{cases}$$

where $x = (x_1, \dots, x_s)$, $y = (y_1, \dots, y_s)$, $C = C_1 \times \dots \times C_s$, $D = D_1 \times \dots \times D_s$.

With the aid of (34.3) and the following Lemma 3.3, it can easily be verified that the function (K_1, \dots, K_s) is concave-convex with effective domain $C \times D$.

Such a saddle function is called separable. In Theorem 3.4 we establish some useful technical facts about separable saddle functions. The proofs of these facts rely on the following similar facts about separable convex functions.

LEMMA 3.3. For $i = 1, \dots, s$ let f_i be a proper convex function on \mathbb{R}^{n_i} with effective domain C_i . Define $C = C_1 \times \dots \times C_s$ and $f(x_1, \dots, x_s) = f_1(x_1) + \dots + f_s(x_s)$. Then the following statements hold:

- (a) f is proper convex with effective domain C ;
- (b) $(cl\ f)(x_1, \dots, x_s) = (cl\ f_1)(x_1) + \dots + (cl\ f_s)(x_s)$;
- (c) f is polyhedral if each f_i is;
- (d) $f^\circ(x_1^\circ, \dots, x_s^\circ) = f_1^\circ(x_1^\circ) + \dots + f_s^\circ(x_s^\circ)$;
- (e) $\partial f(x_1, \dots, x_s) = \partial f_1(x_1) \times \dots \times \partial f_s(x_s)$;
- (f) $(rec\ f)(y_1, \dots, y_s) = (rec\ f_1)(y_1) + \dots + (rec\ f_s)(y_s)$.

PROOF. Assertions (a) and (d) are trivial. Assertion (f) follows immediately from (a) and (8.5). To see (b), let $x = (x_1, \dots, x_s) \in cl\ C = cl\ C_1 \times \dots \times cl\ C_s$ be given and fix any $x_0 = (x_1^0, \dots, x_s^0) \in ri\ C = ri\ C_1 \times \dots \times ri\ C_s$. Define $x_\lambda = (x_1^\lambda, \dots, x_s^\lambda)$ by $x_\lambda = (1 - \lambda)x_0 + \lambda x$ for $0 \leq \lambda < 1$. Then (a) and (7.5) imply that

$$(cl\ f)(x) = \lim_{\lambda \uparrow 1} f(x_\lambda) = \sum \lim_{\lambda \uparrow 1} f_i(x_i^\lambda) = \sum (cl\ f_i)(x_i).$$

On the other hand, if $x \notin cl\ C$ then $x_j \notin cl\ C_j$ for some $1 \leq j \leq s$ and hence (a) and (7.4) imply that $(cl\ f)(x) = +\infty = (cl\ f_j)(x_j) \leq \sum (cl\ f_i)(x_i)$. This proves

(b). To see (c), define $h_i(x_1, \dots, x_s) = f_i(x_i)$ for each i . Then

$$\text{epi } h_i = \{(x_1, \dots, x_s, \mu) \mid (x_i, \mu) \in \text{epi } f_i\}$$

and $\text{epi } f_i$ polyhedral imply that $\text{epi } h_i$ is polyhedral for each i . Hence

(19.4) implies that $f = h_1 + \dots + h_s$ is polyhedral. Finally, we prove (e).

Suppose first that $x = (x_1, \dots, x_s) \notin C$. Then (a) and (23.4) imply that

$$\partial f(x) = \emptyset \text{ and also } \partial f_i(x_i) = \emptyset \text{ for some } 1 \leq i \leq s. \text{ Thus } \partial f_1(x_1) + \dots + \partial f_s(x_s) = \emptyset.$$

Now suppose that $x \in C$. Using (6.1) and (7.5), one can easily verify that,

for a convex function h on \mathbb{R}^n and a subset C of \mathbb{R}^n containing $\text{ri}(\text{dom } h)$,

$x^* \in \partial h(x)$ if and only if

$$h(y) \geq h(x) + \langle x^*, y - x \rangle, \quad \forall y \in C.$$

Applied to the situation at hand, this implies that $x^* = (x_1^*, \dots, x_s^*) \in \partial f(x)$

if and only if

$$\sum f_i(y_i) \geq \sum (f_i(x_i) + \langle x_i^*, y_i - x_i \rangle) \quad (1)$$

for every $(y_1, \dots, y_s) \in C$. Let j be any fixed index. By letting y_i vary over C_i and requiring $y_i = x_i$ for $i \neq j$, (1) implies

$$f_j(y_j) + \sum_{i \neq j} f_i(x_i) \geq f_j(x_j) + \langle x_j^*, y_j - x_j \rangle + \sum_{i \neq j} (f_i(x_i) + \langle x_i^*, x_i - x_i \rangle).$$

Since all the numbers $f_i(x_i)$ are finite, this reduces to

$$f_j(y_j) \geq f_j(x_j) + \langle x_j^*, y_j - x_j \rangle, \quad \forall y_j \in C_j. \quad (2)$$

But this is equivalent to $x_j^* \in \partial f_j(x_j)$. Thus we have shown that (1) implies

$x_j^* \in \partial f_j(x_j)$ for $j = 1, \dots, s$. The converse follows easily by summing the

s inequalities of the form (2). This completes the proof of (e).

THEOREM 3.4. For $i = 1, \dots, s$ let K_i be a closed proper concave-convex function on $\mathbb{R}^m \times \mathbb{R}^n$ with effective domain $C_i \times D_i$. Put $K = (K_1, \dots, K_s)$ and write $C = C_1 \times \dots \times C_s$, $D = D_1 \times \dots \times D_s$, $x = (x_1, \dots, x_s)$ and $y = (y_1, \dots, y_s)$. Then the following statements hold:

- (a) K is closed proper concave-convex with effective domain $C \times D$.
 (b) If $\tilde{K}_i \in [K_i]$ for $i = 1, \dots, s$, then $(\tilde{K}_1, \dots, \tilde{K}_s) \in [K]$ (That is, $[K]$ depends only on $[K_1], \dots, [K_s]$).

(c) The least and greatest members of $[K]$ are given by

$$\underline{K}(x, y) = \begin{cases} \sum K_i(x_i, y_i) & \text{if } x \in C \text{ and } y \in \text{cl } D \\ +\infty & \text{if } x \in C \text{ and } y \notin \text{cl } D \\ -\infty & \text{if } x \notin C \end{cases}$$

and

$$\bar{K}(x, y) = \begin{cases} \sum \bar{K}_i(x_i, y_i) & \text{if } x \in \text{cl } C \text{ and } y \in D \\ -\infty & \text{if } x \notin \text{cl } C \text{ and } y \in D \\ +\infty & \text{if } y \notin D \end{cases}$$

(d) For $j = 1$ and 2 and $(x, y) \in C \times D$,

$$\partial_j K(x, y) = \partial_j K_1(x_1, y_1) \times \dots \times \partial_j K_s(x_s, y_s)$$

(and $\partial K(x, y) = \emptyset$ whenever $(x, y) \notin C \times D$).

- (e) $(K_1^\circ, \dots, K_s^\circ) \in [K^\circ]$
 (f) $(\text{rec}_1 K)(x) = \sum (\text{rec}_1 K_i)(x_i)$ and $(\text{rec}_2 K)(y) = \sum (\text{rec}_2 K_i)(y_i)$
 (g) If f (resp. f_i) denotes the convex parent of K (resp. K_i), then $f(x, y^\circ) = \sum f_i(x_i, y_i^\circ)$. Similarly for concave parents.

(h) If each K_i is polyhedral, then K is polyhedral.

PROOF. (a) It suffices to check that K satisfies the six conditions of (34.3). Let $x \in C$. Then $x_i \in C_i$ together with (34.3) applied to K_i imply that $K_i(x_i, \cdot)$ is a proper convex function with effective domain containing D_i . Since $K(x, y) = \sum K_i(x_i, y_i) + \delta(y|D)$, it follows from Lemma 3.3(a) and (5.2) that $K(x, \cdot)$ is proper convex with effective domain D . Now suppose $x \in \text{ri } C$. Then $x_i \in \text{ri } C_i$, so that (34.3) implies $K(x_i, \cdot)$ is closed and its effective domain actually equals D_i . Thus $K(x, y) = \sum K_i(x_i, y_i)$, and Lemma 3.3(b) implies $K(x, \cdot)$ is closed. This establishes the first two conditions of (34.3) for K . Of the remaining four conditions, two have parallel proofs and the other two are satisfied trivially.

(b) Let $\tilde{K}_i \in [K_i]$ for $i = 1, \dots, s$ and write $\tilde{K} = (\tilde{K}_1, \dots, \tilde{K}_s)$. Since by (34.4) two closed proper saddle functions are equivalent if and only if they have the same kernel, $\tilde{K}_i(x_i, y_i) = K_i(x_i, y_i)$ whenever $(x_i, y_i) \in \text{ri } C_i \times \text{ri } D_i$. Hence \tilde{K} and K agree on $\text{ri } C \times \text{ri } D$. Since equivalent saddle functions have the same effective domain, $\text{dom } \tilde{K}_i = C_i \times D_i$. This implies $\text{dom } \tilde{K} = C \times D$. Therefore \tilde{K} and K have the same kernel.

(c) Since K is closed, Theorem 0.1(b) implies $\underline{K} = \text{cl}_2 K$ and $\bar{K} = \text{cl}_1 K$. If $y \notin D$, then $K(\cdot, y)$ equals ∞ on $\text{ri } C$ and hence $\bar{K}(\cdot, y) = \text{cl}(K(\cdot, y)) = \infty$. Now suppose $y \in D$. As in the proof of part (a), $K(x, y) = \sum K_i(x_i, y_i) - \delta(x|C)$ is proper concave with effective domain C . Since $g(x) = \sum K_i(x_i, y_i)$ is proper concave with $C \subset \text{dom } g \subset \text{cl } C$ by (34.3) and Lemma 3.3(a), it follows from (6.3.1) and (7.3.4) that $(\text{cl}_1 K)(x, y) = (\text{cl } g)(x)$. But $(\text{cl } g)(x) = \sum (\text{cl}_1 K_i)(x_i, y_i)$

by Lemma 3.3(b). Since $\text{cl}_1 K_1 = \bar{K}_1$, this establishes the formula for \bar{K} . The other formula is proved similarly.

(d) By part (a) and (37.4), $\text{dom } \partial K \subset C \times D$. Suppose $(x, y) \in C \times D$. By (37.4.1), $\partial K(x, y) = \partial \bar{K}(\cdot, y)(x) \times \partial \underline{K}(x, \cdot)(y)$. But from part (c) and Lemma 3.3(e), $\partial \bar{K}(\cdot, y)(x) = \partial \bar{K}_1(\cdot, y_1)(x_1) \times \dots \times \partial \bar{K}_s(\cdot, y_s)(x_s)$ where by (37.4.1) the \bar{K}_1 can be replaced by K_1 . This establishes the assertion for $j = 1$, and the case $j = 2$ is exactly the same.

(e) The proof is by induction. First observe that separable saddle functions can be given an equivalent, inductive definition. Namely, for $s = 2$ let the definition be as given above, and for $s = 2$ let the definition be as given above, and for $s > 2$ define $(K_1, \dots, K_s) = ((K_1, \dots, K_{s-1}), K_s)$ where a space of the form $(R^{m_1} \times \dots \times R^{m_{s-1}}) \times R^{m_s}$ is identified with $R^{m_1} \times \dots \times R^{m_s}$. For the purpose of this proof we adopt this inductive definition. Suppose the assertion has already been proved for the case $s = 2$, and let $s > 2$ be fixed. Since $(K_1^\circ, \dots, K_s^\circ) = ((K_1^\circ, \dots, K_{s-1}^\circ), K_s^\circ)$ by definition, the inductive hypothesis

$$(K_1^\circ, \dots, K_{s-1}^\circ) \in ((K_1, \dots, K_{s-1})^\circ)$$

together with parts (a) and (b) imply that $(K_1^\circ, \dots, K_s^\circ)$ is equivalent to $((K_1, \dots, K_{s-1})^\circ, K_s^\circ)$. But by the case $s = 2$ this is contained in $((K_1, \dots, K_{s-1}), K_s)^\circ$, which by definition is the same as $(K_1, \dots, K_s)^\circ$. Thus, part (e) will be proved once the case $s = 2$ is established. So let $s = 2$ and write $\text{dom } K_1^\circ = C_1^\circ \times D_1^\circ$. By (36.3) and (36.1),

$$\begin{aligned}
\underline{K}^{\circ}(x^{\circ}, y^{\circ}) &= \sup_{y \in D} \inf_{x \in C} (\Sigma \langle x_1, x_1^{\circ} \rangle + \langle y_1, y_1^{\circ} \rangle \cdot K_1(x_1, y_1)) \\
&\leq \sup_{y_2 \in D_2} \inf_{x_2 \in C_2} (\langle x_2, x_2^{\circ} \rangle + \langle y_2, y_2^{\circ} \rangle \cdot K_2(x_2, y_2) + \underline{K}_1^{\circ}(x_1^{\circ}, y_1^{\circ})) \\
&= \begin{cases} \Sigma \underline{K}_1^{\circ}(x_1^{\circ}, y_1^{\circ}) & \text{if } x_1^{\circ} \in C_1^{\circ} \text{ and } y_1^{\circ} \in \text{dom } \underline{K}_1^{\circ}(x_1^{\circ}, \cdot) \\ +\infty & \text{if } x_1^{\circ} \in C_1^{\circ} \text{ and } y_1^{\circ} \notin \text{dom } \underline{K}_1^{\circ}(x_1^{\circ}, \cdot) \\ -\infty & \text{if } x_1^{\circ} \notin C_1^{\circ} \end{cases}
\end{aligned}$$

Moreover, in the event that $x_1^{\circ} \in C_1^{\circ}$ and $y_1^{\circ} \in \text{dom } \underline{K}_1^{\circ}(x_1^{\circ}, \cdot)$ we have

$$\Sigma \underline{K}_1^{\circ}(x_1^{\circ}, y_1^{\circ}) = \begin{cases} \Sigma \underline{K}_1^{\circ}(x_1^{\circ}, y_1^{\circ}) \in \mathbb{R} & \text{if } x_2^{\circ} \in C_2^{\circ} \text{ and } y_2^{\circ} \in \text{dom } \underline{K}_2^{\circ}(x_2^{\circ}, \cdot) \\ +\infty & \text{if } x_2^{\circ} \in C_2^{\circ} \text{ and } y_2^{\circ} \notin \text{dom } \underline{K}_2^{\circ}(x_2^{\circ}, \cdot) \\ -\infty & \text{if } x_2^{\circ} \notin C_2^{\circ} \end{cases}$$

Also, $x_1^{\circ} \in C_1^{\circ}$ implies $D_1^{\circ} \subset \text{dom } \underline{K}_1^{\circ}(x_1^{\circ}, \cdot)$ by (34.3). If $C^{\circ} = C_1^{\circ} \times C_2^{\circ}$ and $D^{\circ} = D_1^{\circ} \times D_2^{\circ}$, then the above facts imply

$$\text{dom}_1 \underline{K}^{\circ} \subset C^{\circ}, D^{\circ} \subset \text{dom}_2 \underline{K}^{\circ},$$

and

$$\underline{K}^{\circ}(x^{\circ}, y^{\circ}) \leq \Sigma \underline{K}_1^{\circ}(x_1^{\circ}, y_1^{\circ}) \text{ whenever } x^{\circ} \in C^{\circ} \text{ or } y^{\circ} \in D^{\circ}.$$

Parallel reasoning starting from $\bar{K}^{\circ}(x^{\circ}, y^{\circ})$ yields that

$$C^{\circ} \subset \text{dom}_1 \bar{K}^{\circ}, \text{dom}_2 \bar{K}^{\circ} \subset D^{\circ},$$

and

$$\Sigma \bar{K}_1^{\circ}(x_1^{\circ}, y_1^{\circ}) \leq \bar{K}^{\circ}(x^{\circ}, y^{\circ}) \text{ whenever } x^{\circ} \in C^{\circ} \text{ or } y^{\circ} \in D^{\circ}.$$

Therefore $\text{dom } K^{\circ} = C^{\circ} \times D^{\circ}$ and

$$\underline{K}^\circ(x^\circ, y^\circ) \leq \sum K_1^\circ(x_1^\circ, y_1^\circ) \leq \bar{K}^\circ(x^\circ, y^\circ)$$

whenever $x^\circ \in C^\circ$ or $y^\circ \in D^\circ$. By applying Theorem 0.1(b) to K° , it follows that (K_1°, K_2°) and K° have the same kernel. Since they are both closed and proper, (34.4) implies they are equivalent.

(f) From the definitions, Theorem 3.4(a), (34.3) and Lemma 3.3(f) it follows that

$$\begin{aligned} (\text{rec}_2 K)(y) &= \sup(\text{rec } K(x, \cdot))(y) | x \in \cap C) \\ &= \sup(\sum (\text{rec } K_1(x_1, \cdot))(y_1) | x_1 \in \cap C_1, \dots, x_s \in \cap C_s) \\ &= \sum \sup((\text{rec } K_1(x_1, \cdot))(y_1) | x_1 \in \cap C_1) \\ &= \sum (\text{rec}_2 K_1)(y_1). \end{aligned}$$

The other formula is proved similarly.

(g) By part (c), $\underline{K}(x, y) = \sum \underline{K}_1(x_1, y_1)$ whenever $x \in C$. Hence Theorem 0.1(a) implies that $f(x, y^\circ) = \sup(\langle y, y^\circ \rangle - \sum \underline{K}_1(x_1, y_1)) = \sum \sup(\langle y_1, y_1^\circ \rangle - \underline{K}_1(x_1, y_1)) = \sum f_1(x_1, y_1^\circ)$ whenever $x \in C$. On the other hand, if $x \notin C$ then Theorem 0.1(a) implies that $f(x, \cdot)$ and $f_1(x_1, \cdot)$ for some $1 \leq j \leq s$ are constantly $-\infty$. Since each K_1 is proper, each f_1 is proper and hence never $-\infty$. Therefore $(x, y^\circ) \rightarrow \sum f_1(x_1, y_1^\circ)$ is constantly $-\infty$ whenever $x \notin C$. This establishes the formula.

(h) By part (g) and Lemma 3.3(c).

For the remainder of §3 let certain notation remain fixed as follows. For $i = 1, \dots, s$ let K_i be a closed proper concave-convex function on $\mathbb{R}^m \times \mathbb{R}^n$ with effective domain $C_i \times D_i$. Write $C = C_1 \cap \dots \cap C_s$ and $D = D_1 \cap \dots \cap D_s$, and define $K = (K_1, \dots, K_s)$. Let A_j map each $x \in \mathbb{R}^m$ into the s -tuple

(x, \dots, x) , let A_2 map each $y \in R^n$ into the s -tuple (y, \dots, y) , and put $A = A_1 \times A_2$.

In what follows we shall frequently use the condition

$$ri(\text{dom } K_1) \cap \dots \cap ri(\text{dom } K_s) \neq \emptyset. \quad (*)$$

The next lemma dualizes it.

LEMMA 3.5. The condition $ri(\text{dom}_1 K_1) \cap \dots \cap ri(\text{dom}_1 K_s) \neq \emptyset$ is equivalent to

$$\sum x_i^\circ = 0 \text{ and } \sum (rec_1 K_1)(x_i^\circ) \geq 0 \implies \sum (rec_1 K_1)(-x_i^\circ) \geq 0.$$

Similarly, the condition $ri(\text{dom}_2 K_1) \cap \dots \cap ri(\text{dom}_2 K_s) \neq \emptyset$ is equivalent to

$$\sum y_i^\circ = 0 \text{ and } \sum (rec_2 K_1)(y_i^\circ) \leq 0 \implies \sum (rec_2 K_1)(-y_i^\circ) \leq 0.$$

PROOF. Apply Lemma 0.7 to the saddle function $(K_1^\circ, \dots, K_s^\circ)$ and the subspace $\{(x_1^\circ, \dots, x_s^\circ) \mid x_1^\circ = \dots = x_s^\circ\}$ and simplify using Theorem 3.4. The second assertion is proved similarly.

The next theorem (together with Theorem 3.4) enables us to apply the results of §§1 and 2 to the equivalence class $[K_1 + \dots + K_s]$.

THEOREM 3.6. Assume $(*)$. Then $[K_1 + \dots + K_s]$ is defined and equals $[KA]$.

PROOF. Theorem 3.2 implies $[K_1 + \dots + K_s]$ is defined and has as its kernel the function $(x, y) \rightarrow K_1(x, y) + \dots + K_s(x, y)$ on $ri C \times ri D$. Theorems 3.4(a) and 1.2 imply that $[KA]$ exists, and it is easy to check that its kernel is the function just given. The theorem now follows from (3.4).

COROLLARY 3.6.1. Assume $(*)$. If each $[K_i]$ is polyhedral, then $[K_1 + \dots + K_s]$ is polyhedral.

PROOF. By Theorem 3.4(h) and Corollary 1.4.1.

COROLLARY 3.6.2. Assume (*). If h (resp. k) denotes the convex (resp. concave) parent of $[K_1 + \dots + K_g]$ and f_1 (resp. g_1) denotes the convex (resp. concave) parent of $[K_1]$, then

$$h(x, y^\circ) = \inf(\sum f_1(x, y_1^\circ) \mid \sum y_1^\circ = y^\circ)$$

and

$$k(x^\circ, y) = \sup(\sum g_1(x_1^\circ, y) \mid \sum x_1^\circ = x^\circ) .$$

PROOF. By Theorems 1.4 and 3.4(g) .

COROLLARY 3.6.3. Assume (*). Then

$$\text{dom}(K_1 + \dots + K_g)^\circ \subset \text{dom } K_1^\circ + \dots + \text{dom } K_g^\circ .$$

In particular, if f_1 and g_1 are as in Corollary 3.6.2, then

$$\text{ri}(\text{dom}_1(K_1 + \dots + K_g)^\circ) = \cup \{ \sum \text{ri}(\text{dom } g_1(\cdot, y)) \mid y \in \text{ri } D \}$$

and

$$\text{ri}(\text{dom}_2(K_1 + \dots + K_g)^\circ) = \cup \{ \sum \text{ri}(\text{dom } f_1(x, \cdot)) \mid x \in \text{ri } C \}$$

where these formulas also hold with "ri" deleted throughout.

PROOF. Apply Corollary 1.4.2, using Theorem 3.4(g) and Lemma 3.3(a) to simplify.

Convex function theory has a result corresponding to the inclusion in Corollary 3.6.3. It is that

$$\text{dom}(f_1 + \dots + f_g)^\circ = \text{dom } f_1^\circ + \dots + \text{dom } f_g^\circ$$

whenever f_1, \dots, f_g are proper convex functions satisfying $\text{ri}(\text{dom } f_1) \cap \dots \cap \text{ri}(\text{dom } f_g) \neq \emptyset$ (see(16.4)). One might hope in the saddle function case to have at least the inclusions

$$ri(\text{dom } K_1^\circ + \dots + \text{dom } K_s^\circ) \subset \text{dom}(K_1 + \dots + K_s)^\circ \subset \text{dom } K_1^\circ + \dots + \text{dom } K_s^\circ$$

satisfied whenever K_1, \dots, K_s are closed proper concave-convex functions satisfying $ri(\text{dom } K_1) \cap \dots \cap ri(\text{dom } K_s) \neq \emptyset$. However this can fail drastically, as can be seen by taking $s=2$ and putting $K_1(x, y) = \langle x, y \rangle$ and $K_2(x, y) = -\langle x, y \rangle$ on $\mathbb{R}^m \times \mathbb{R}^m$. In this case $\text{dom}(K_1 + K_2)^\circ = \{0\} \times \{0\}$ whereas $\text{dom } K_1^\circ = \mathbb{R}^m \times \mathbb{R}^m = \text{dom } K_2^\circ$. Lemma 3.7 and Theorem 3.11 give conditions which guarantee that such "collapsing" $\text{dom}(K_1 + \dots + K_s)^\circ$ cannot occur.

LEMMA 3.7. Assume (*). Then for $j = 1$ and 2 ,

$$cl(\text{dom}_j(K_1 + \dots + K_s)^\circ) = cl(\text{dom}_j K_1^\circ + \dots + \text{dom}_j K_s^\circ)$$

if and only if

$$rec_j(K_1 + \dots + K_s) = rec_j K_1 + \dots + rec_j K_s.$$

PROOF. By Theorems 3.6 and Lemma 1.5.

LEMMA 3.8. Assume (*). Then

$$(rec_1(K_1 + \dots + K_s))(x) = \inf \{ \sum (rec K_i(\cdot, y))(x) \mid y \in ri D \}$$

and

$$(rec_2(K_1 + \dots + K_s))(y) = \sup \{ \sum (rec K_i(x, \cdot))(y) \mid x \in ri C \}.$$

PROOF. By Theorem 3.6 the formulas in Lemma 1.6 can be applied. Simplify using (34.3) and Lemma 3.3(f).

The next theorem parallels the result obtained by Rockafellar [42], Moreau [36], and others for the subdifferential of a sum of convex functions.

THEOREM 3.9. Assume (*) . Then

$$\partial(K_1 + \dots + K_g)(x, y) = \partial K_1(x, y) + \dots + \partial K_g(x, y)$$

for each $(x, y) \in R^m \times R^n$.

PROOF. Since $\text{dom}(K_1 + \dots + K_g) = C \times D = \text{dom } K_1 \cap \dots \cap \text{dom } K_g$,
(37.4) implies that $\partial(K_1 + \dots + K_g)(x, y)$ and $\partial K_j(x, y)$ are empty (for some j)
whenever $(x, y) \notin C \times D$. So suppose that $(x, y) \in C \times D$. Then Theorems 3.6
and 1.3 imply that

$$\partial(K_1 + \dots + K_g)(x, y) = A^* \partial K(A(x, y)) .$$

The formula follows from this together with Theorem 3.4(d) and the definitions,
after observing that A_1^* and A_2^* are just the appropriate addition linear
transformations.

The next theorem identifies certain members of the equivalence class
conjugate to $[K_1 + \dots + K_g]$.

THEOREM 3.10. Assume (*) . Let $\text{dom } K_1^* = C_1^* \times D_1^*$, and define func-
tions ϕ and ψ on $R^m \times R^n$ by

$$\phi(z, w) = \sup_{\substack{\sum z_i = z \\ z_i \in C_1^*}} \inf_{\sum w_i = w} \sum K_1^*(z_i, w_i)$$

and

$$\psi(z, w) = \inf_{\substack{\sum w_i = w \\ w_i \in D_1^*}} \sup_{\sum z_i = z} \sum \bar{K}_1^*(z_i, w_i) .$$

Then $[(K_1 + \dots + K_g)^*]$ contains each concave-convex function J on $R^m \times R^n$

satisfying $\phi \leq J \leq \psi$. If each K_1 is polyhedral, then $[(K_1 + \dots + K_s)^*]$ is polyhedral.

PROOF. By Theorems 3.6 and 1.8, $[(K_1 + \dots + K_s)^*]$ contains each saddle function lying between two certain functions J_1 and J_2 . By parts (e) and (c) of Theorem 3.4, one can easily show that $J_1 = \phi$ and $J_2 = \psi$. The polyhedral assertion is immediate from Corollary 3.6.1 and the fact that $[K^*]$ is polyhedral whenever $[K]$ is.

The fact that $[(K_1 + \dots + K_s)^*]$ contains ϕ and ψ suggests writing

$$[(K_1 + \dots + K_s)^*] = [K_1^*] \square \dots \square [K_s^*]$$

and calling this class the minimax convolution of $[K_1^*], \dots, [K_s^*]$. This is the saddle function analogue of the operation of infimal convolution on convex functions. The identity above expresses the fact that the operations of addition and minimax convolution of equivalence classes are dual with respect to conjugacy, or in other words, that the conjugate of the sum of equivalence classes is the minimax convolute of the conjugates.

The next theorem gives information concerning attainment of the extrema appearing in the definitions of ϕ and ψ .

THEOREM 3.11. Let ϕ and ψ be defined as in Theorem 3.10, and assume that whenever $(z_1, w_1) \in \text{ri}(\text{dom } K_1^*)$ for $1 = 1, \dots, s$ the following conditions are satisfied:

- (a) $\Sigma \bar{w}_1 = 0$ and $\Sigma (\text{rec } K_1^*(z_1, \cdot))(\bar{w}_1) \leq 0 \implies \Sigma (\text{rec } K_1^*(z_1, \cdot))(-\bar{w}_1) \leq 0$
- (b) $\Sigma \bar{z}_1 = 0$ and $\Sigma (\text{rec } K_1^*(\cdot, w_1))(\bar{z}_1) \geq 0 \implies \Sigma (\text{rec } K_1^*(\cdot, w_1))(-\bar{z}_1) \geq 0$.

Then the conclusions of Theorem 3.10 hold and

$$\text{ri}(\text{dom } K_1^* + \dots + \text{dom } K_s^*) \subset \text{dom}(K_1 + \dots + K_s)^* \subset \text{dom } K_1^* + \dots + \text{dom } K_s^* .$$

Moreover, for each $(z, w) \in \text{ri}(\text{dom } K_1^* + \dots + \text{dom } K_s^*)$ there exist nonempty closed convex sets $Z \subset \mathbb{R}^{sm}$ and $W \subset \mathbb{R}^{sn}$ such that for each $(\bar{z}_1, \dots, \bar{z}_s) \in Z$ and $(\bar{w}_1, \dots, \bar{w}_s) \in W$ the following statements hold:

- (i) $\Sigma(\bar{z}_1, \bar{w}_1) = (z, w)$ and $(\bar{z}_1, \bar{w}_1) \in \text{dom } \partial K_1^*$ for each i ;
- (ii) $\phi(z, w) = \Sigma K_1^*(\bar{z}_1, \bar{w}_1) = \psi(z, w)$;
- (iii) $\Sigma K_1^*(z_1, \bar{w}_1) \leq \Sigma K_1^*(\bar{z}_1, \bar{w}_1) \leq \Sigma K_1^*(\bar{z}_1, w_1)$ whenever $\Sigma(z_1, w_1) = (z, w)$ and $(z_1, w_1) \in \text{dom } K_1^*$ for each i .

PROOF. By parts (e) and (c) of Theorem 3.4 together with Lemma 3.3(f), A^* and $K^* = (K_1^*, \dots, K_s^*)$ satisfy the hypotheses of Theorem 2.4. The assertions are immediate from this and Theorem 3.4(d).

If conditions (a) and (b) above are actually satisfied whenever $(z_1, w_1) \in \text{dom } K_1^*$ for $i = 1, \dots, s$, then Theorem 2.5 implies that $\text{dom}(K_1 + \dots + K_s)^* = \text{dom } K_1^* + \dots + \text{dom } K_s^*$ and moreover that ϕ and ψ are respectively the least and greatest members of $[(K_1 + \dots + K_s)^*]$.

The following lemma may be useful in applying Theorem 3.11 and the above remark. Notice, for example, that the conditions given are satisfied when each of the sets $\text{dom } K_i^*$ is bounded.

LEMMA 3.12. The following condition implies that condition (a) of Theorem 3.11 is satisfied for each choice of $z_1 \in \text{dom}_1 K_1^*, \dots, z_s \in \text{dom}_1 K_s^*$:

- (c) $\Sigma w_1 = 0$ and $w_1 \in 0^+ \text{cl}(\text{dom}_2 K_1^*)$ for each i imply that $w_1 = 0$ for each i .

Similarly, the following condition implies that condition (b) of Theorem 3.11 is satisfied for each choice of $w_1 \in \text{dom}_2 K_1^*, \dots, w_s \in \text{dom}_2 K_s^*$:

- (d) $\sum z_i = 0$ and $z_i \in 0^+ \text{cl}(\text{dom}_1 K_i^*)$ for each i imply that $z_i = 0$ for each i .

PROOF. Apply Lemma 2.6 to A^* and $K^* = (K_1^*, \dots, K_s^*)$. Condition (c) (resp. (d)) corresponds to condition (c_1) (resp. (d_1)) of Lemma 2.6, and condition (a) (resp. (b)) of Theorem 3.11 corresponds to condition (a_3) (resp. (b_3)) of Lemma 2.3.

Conditions (c) and (d) of Lemma 3.12 may be given alternate characterizations with the aid of the next lemma.

LEMMA 3.13. Let P_1, \dots, P_s be convex cones in \mathbb{R}^n which contain the origin. Then the following conditions are equivalent:

- (i) $\sum p_i = 0$ and $p_i \in P_i$ for each i imply $p_i = 0$ for each i ;
 (ii) $(-P_j) \cap (\text{conv} \bigcup_{i \neq j} P_i) = \{0\}$ for each $j = 1, \dots, s$.

PROOF. First, observe that for each j , (3.3) implies

$$\text{conv} \bigcup_{i \neq j} P_i = \bigcup_{i \neq j} \{ \sum \lambda_i P_i \mid 0 \leq \lambda_i \text{ and } \sum \lambda_i = 1 \}.$$

From this it follows that $\text{conv} \bigcup_{i \neq j} P_i = \sum_{i \neq j} P_i$. Thus, condition (ii) fails if and only if

$$\exists j \text{ and } \exists p_j \in P_j \text{ such that } 0 \neq -p_j \in \sum_{i \neq j} P_i.$$

This occurs if and only if

$$\exists p_1 \in P_1, \dots, \exists p_s \in P_s \text{ and } \exists j \text{ such that } 0 \neq -p_j = \sum_{i \neq j} p_i,$$

which occurs if and only if condition (1) fails.

We conclude this section with an example concerning maximal monotone operators arising from saddle functions. This will suggest a conjecture about arbitrary maximal monotone operators.

By (37.5.2) (see also [49]), each closed proper concave-convex function K on $R^m \times R^n$ induces a maximal monotone operator T (generally multivalued) from $R^m \times R^n$ to $R^m \times R^n$ by means of the formula

$$T(x, y) = \{(-x^*, y^*) \mid (x^*, y^*) \in \partial K(x, y)\}.$$

By (37.4.1), T depends only on the equivalence class containing K . If $R(\cdot)$ denotes the range of an operator and B is the linear transformation which sends (x^*, y^*) to $(-x^*, y^*)$, then (37.5) implies that

$$R(T) = B \operatorname{dom} \partial K^*$$

whenever T arises from K as above.

EXAMPLE 3.14. Assume that conditions (c) and (d) of Lemma 3.12 are satisfied. Then the hypotheses of Theorem 3.11 are met, and these in turn imply that condition (*) is satisfied. Let B be the linear transformation defined above, let T_i be the maximal monotone operator induced by K_i as described above, and similarly (using Theorem 3.6) let T be the maximal monotone operator induced by $\sum K_i$. By (37.4), (6.3.1) and (9.1) it follows that $\operatorname{cl} R(T_i) = B \operatorname{cl} \operatorname{dom} K_i^*$, and similarly $\operatorname{cl} R(T) = B \operatorname{cl} \operatorname{dom} (\sum K_i)^*$. Theorem 3.11 and (6.3.1) imply $\operatorname{cl} \operatorname{dom} (\sum K_i)^* = \operatorname{cl} \sum \operatorname{dom} K_i^*$. Combining these facts with (6.6.2) yields $\sum \operatorname{cl} R(T_i) \subset \operatorname{cl} R(T)$. Since Theorem 3.9 implies $\sum T_i = T$, this shows that $\sum T_i$ is a maximal monotone operator satisfying

$$\sum \operatorname{cl} R(T_i) \subset \operatorname{cl} R(\sum T_i). \quad (1)$$

Furthermore, it can be deduced from (1), $R(\sum T_i) \subset \sum R(T_i)$ and (8.1) that

$$\sum 0^+ \text{cl} R(T_i) \subset 0^+ \text{cl} R(\sum T_i).$$

It is easy to show that conditions (c) and (d) of Lemma 3.12 can be reformulated equivalently as follows:

$$\begin{aligned} \sum z_i = 0 \text{ and } z_i \in 0^+ \text{cl} R(T_i) \\ \text{for each } i \text{ implies } z_i = 0 \text{ for each } i \end{aligned} \quad (\dagger\dagger)$$

Now $(\dagger\dagger)$ and (9.1.1) imply $\sum \text{cl} R(T_i) = \text{cl} \sum R(T_i)$, and $R(\sum T_i) \subset \sum R(T_i) \subset \sum \text{cl} R(T_i)$ holds trivially. Combining these facts with (1) yields

$$\sum \text{cl} R(T_i) = \text{cl} R(\sum T_i). \quad (2)$$

Furthermore, from (2), $(\dagger\dagger)$ and (9.1.1) it follows immediately that

$$\sum 0^+ \text{cl} R(T_i) = 0^+ \text{cl} R(\sum T_i).$$

It is known that $\sum T_i$ is a maximal monotone operator satisfying (1) whenever each T_i is the subdifferential of a closed proper convex function on R^n and the condition

$$\text{ri} \, \partial(T_1) \cap \dots \cap \text{ri} \, \partial(T_s) \neq \emptyset \quad (\dagger)$$

is satisfied, where $\partial(T) = \{z \mid T(z) \neq \emptyset\}$. Moreover, in this situation (2) actually holds if $(\dagger\dagger)$ is satisfied.

On the other hand, (1) fails in general for maximal monotone operators satisfying (\dagger) . (For example, take $s = 2$ and consider the T_i 's induced by the saddle functions K_1 and K_2 defined following Corollary 3.6.3.) It is not known, though, whether (2) holds for arbitrary maximal monotone operators satisfying $(\dagger\dagger)$. But the fact that this formula does hold for those operators

arising from saddle functions leads one to conjecture that it holds in general.

This is because such operators, unlike the subdifferentials of convex functions, exhibit most of the pathology of arbitrary maximal monotone operators. Indeed, this last fact is one of the main motivations for studying saddle functions.

§4. The Partial Conjugacy Operation

In this short section the results of §1 are used to develop another operation on equivalence classes of closed proper saddle functions. By virtue of its similarity to the basic conjugacy operation, this is called the partial conjugacy operation. It follows from Theorems 4.1 and 4.2 that the partial conjugacy operation induces another symmetric one-to-one correspondence among closed proper equivalence classes. In §5 this correspondence will be used to assign a well-defined Lagrangian to each dual pair of generalized saddle programs. The symmetric one-to-one character of this assignment will be used in §6 to establish the negative result that there exists no good Lagrange multiplier principle for ordinary saddle programs in general.

Throughout §4 let K be a closed proper concave-convex function on $(R^p \times R^m) \times (R^q \times R^n)$, and let W_1 and W_2 be functions on $(R^p \times R^n) \times (R^q \times R^m)$ defined by

$$W_1(u^*, y, v^*, x) = \sup_v \inf_u \{ \langle u^*, u \rangle + \langle v^*, v \rangle - K(u, x, v, y) \}$$

and

$$W_2(u^*, y, v^*, x) = \inf_u \sup_v \{ \langle u^*, u \rangle + \langle v^*, v \rangle - K(u, x, v, y) \} .$$

THEOREM 4.1. The functions W_1 and W_2 belong to an equivalence class $[W]$ of closed proper concave-convex functions. Furthermore, $[W]$ depends only on $[K]$, and $[W]$ is polyhedral if $[K]$ is.

PROOF. Define a linear transformation $A = A_1 \times A_2$ and a function H by

$$A_1(v, u^*, y) = (u^*, y) ,$$

$$A_2(u, v^*, x) = (v^*, x) ,$$

and

$$H(v, u^*, y, u, v^*, x) = \langle u, u^* \rangle + \langle v, v^* \rangle - K(u, x, v, y).$$

Clearly H is closed proper concave-convex on $(R^q \times R^p \times R^n) \times (R^p \times R^q \times R^m)$.

If $(v, y) \in \text{ri}(\text{dom}_2 K)$, (34.3) implies that the function

$$(u, v^*, x) \rightarrow -K(u, x, v, y)$$

is closed proper convex, and by (8.5) its recession function can be shown to be

$$(u, v^*, x) \rightarrow -(\text{rec } K(\cdot, \cdot, v, y))(u, x).$$

Also, the function

$$(u, v^*, x) \rightarrow \langle u, u^* \rangle + \langle v, v^* \rangle$$

is closed proper convex and coincides with its recession function. Hence (9.3)

implies that

$$(\text{rec } H(v, u^*, y, \cdot, \cdot, \cdot))(u, v^*, x) = \langle u, u^* \rangle + \langle v, v^* \rangle - (\text{rec } K(\cdot, \cdot, v, y))(u, x)$$

whenever $(v, y) \in \text{ri}(\text{dom}_2 K)$. Therefore $A_2^{-1}\{0\} \cap (\text{rec cone}_2 H)$ equals

$$\{(u, 0, 0) \mid \langle u, u^* \rangle - (\text{rec } K(\cdot, \cdot, v, y))(u, 0) \leq 0, \forall u^* \in R^p, \forall (v, y) \in \text{ri}(\text{dom}_2 K)\}.$$

Now by (34.3) and (8.5), $(v, y) \in \text{ri}(\text{dom}_2 K)$ implies that $\text{rec } K(\cdot, \cdot, v, y)$ is

never $+\infty$. It follows that $A_2^{-1}\{0\} \cap (\text{rec cone}_2 H)$ is the nullspace of

$R^p \times R^q \times R^m$. Similarly, $A_1^{-1}\{0\} \cap (\text{rec cone}_1 H)$ is the nullspace of $R^q \times R^p \times R^n$.

Therefore by Lemma 1.9, $\text{range } A^* \cap \text{ri}(\text{dom } H^*) \neq \emptyset$. The first two assertions

of the theorem now follow from Theorem 1.8 and the fact that $\underline{K} \leq \tilde{K} \leq \bar{K}$ when-

ever $\tilde{K} \in [K]$. If K is polyhedral, then Corollary 3.6.1 implies H is poly-

hedral and hence Theorem 1.8 implies $[AH] = [W]$ is polyhedral.

The equivalence class $[W]$ containing W_1 and W_2 is called the partial conjugate of $[K]$ in u and v , and the operation which sends $[K]$ to $[W]$ is called partial conjugacy. This terminology is suggested by the fact that forming $[W]$ involves only parts of the arguments of K , whereas forming the (ordinary) conjugate $[K^*]$ involves all of the arguments of K .

THEOREM 4.2. The partial conjugate of $[W]$ in u^* and v^* is $[K]$.

PROOF. By Theorem 4.1, $[W]$ contains the function \tilde{W} , where $\tilde{W}(u^*, y, v^*, x) = \inf_{\tilde{u}} \sup_{\tilde{v}} (\langle \tilde{u}, u^* \rangle + \langle \tilde{v}, v^* \rangle + K(\tilde{u}, x, \tilde{v}, y))$. Hence the partial conjugate of $[W]$ in u^* and v^* contains the function M given by

$$\begin{aligned} M(u, x, v, y) &= \sup_{v^*} \inf_{u^*} (\langle u, u^* \rangle + \langle v, v^* \rangle - \tilde{W}(u^*, y, v^*, x)) \\ &= \sup_{v^*} \inf_{u^*} \sup_{\tilde{u}} \inf_{\tilde{v}} (\langle u^*, u - \tilde{u} \rangle + \langle v^*, v - \tilde{v} \rangle + K(\tilde{u}, x, \tilde{v}, y)) . \end{aligned}$$

By the same technique used in the proof of Theorem 4.1 it can be verified that $\text{range } B^* \cap \text{ri}(\text{dom } J^*) \neq \emptyset$ where $B = B_1 \times B_2$ and J are given by

$$B_1(v^*, \tilde{u}, u, x) = (u, x) ,$$

$$B_2(u^*, \tilde{v}, v, y) = (v, y) ,$$

and

$$J(v^*, \tilde{u}, u, x, u^*, \tilde{v}, v, y) = \langle u^*, u - \tilde{u} \rangle + \langle v^*, v - \tilde{v} \rangle + K(\tilde{u}, x, \tilde{v}, y) .$$

Therefore Theorem 1.8 implies that $[BJ]$ is well-defined. Now by (36.1) and Theorem 0.1(b) it follows easily that M and N belong to $[BJ]$, where N is given by

$$N(u, x, v, y) = \sup_{\tilde{u}} \inf_{\tilde{v}} \sup_{v^*} \inf_{u^*} \{ \langle u^*, u - \tilde{u} \rangle + \langle v^*, v - \tilde{v} \rangle + K(\tilde{u}, x, \tilde{v}, y) \} .$$

Thus, to complete the proof it suffices to show that $N \in [K]$.

Let u, x, v, y be arbitrary but fixed. For each \bar{u} define

$$p(\bar{u}) = \inf_{\bar{v}} \sup_{v^*} \inf_{u^*} \{ \langle u^*, u - \bar{u} \rangle + \langle v^*, v - \bar{v} \rangle + \underline{K}(\bar{u}, x, \bar{v}, y) \}.$$

Observe that

$$N(u, x, v, y) = \sup \{ p(\bar{u}) \mid \bar{u} \in U \}, \quad (1)$$

where $U = \{ \bar{u} \mid (\bar{u}, x) \in \text{dom}_1 K \}$. Indeed, if $(\bar{u}, x) \notin \text{dom}_1 K$ then $\underline{K}(\bar{u}, x, \cdot, \cdot) = -\infty$ so that $p(\bar{u}) = -\infty$. Thus,

$$N(u, x, v, y) = -\infty = \underline{K}(u, x, v, y) \quad (2)$$

whenever $U = \emptyset$. Now assume $U \neq \emptyset$. For each $\bar{u} \in U$, $\underline{K}(\bar{u}, x, \cdot, \cdot)$ is never $-\infty$ and hence

$$p(\bar{u}) = \inf_{\bar{v} \in V(\bar{u})} \sup_{v^*} \inf_{u^*} \{ \langle u^*, u - \bar{u} \rangle + \langle v^*, v - \bar{v} \rangle + \underline{K}(\bar{u}, x, \bar{v}, y) \},$$

where $V(\bar{u}) = \{ \bar{v} \mid \underline{K}(\bar{u}, x, \bar{v}, y) < +\infty \}$. This implies $p(\bar{u}) = +\infty$ whenever $V(\bar{u}) = \emptyset$.

Hence (1) implies $N(u, x, v, y) = +\infty$ if there exists a $\bar{u} \in U$ such that $V(\bar{u}) = \emptyset$.

But for such a \bar{u} , $\underline{K}(\bar{u}, x, v, y) = +\infty$. Therefore

$$N(u, x, v, y) = +\infty = \bar{K}(u, x, v, y) \quad (3)$$

whenever there exists a $\bar{u} \in U$ such that $V(\bar{u}) = \emptyset$. Finally, assume $U \neq \emptyset$

and $V(\bar{u}) \neq \emptyset$ for every $\bar{u} \in U$. Then for each $\bar{u} \in U$,

$$p(\bar{u}) = \inf_{\bar{v} \in V(\bar{u})} \{ \underline{K}(\bar{u}, x, \bar{v}, y) + \sup_{v^*} \{ \langle v^*, v - \bar{v} \rangle + \inf_{u^*} \{ \langle u^*, u - \bar{u} \rangle \} \} \}.$$

Hence $p(\bar{u}) = -\infty$ whenever $u \neq \bar{u} \in U$, while for $u = \bar{u} \in U$ we have that

$$p(u) = \inf_{\bar{v} \in V(u)} \{ \underline{K}(u, x, \bar{v}, y) + \sup_{v^*} \{ \langle v^*, v - \bar{v} \rangle \} \}$$

$$= \inf \{ \underline{K}(u, x, \tilde{v}, y) \mid \tilde{v} \in V(u), \tilde{v} = v \}$$

$$= \begin{cases} +\infty & \text{if } v \notin V(u) \\ \underline{K}(u, x, v, y) & \text{if } v \in V(u) \end{cases}.$$

Hence (1) implies that in this case

$$N(u, x, v, y) = \sup \{ p(\tilde{u}) \mid \tilde{u} \in U, \tilde{u} = u \}$$

$$= \begin{cases} -\infty & \text{if } u \notin U \\ +\infty & \text{if } u \in U \text{ and } v \notin V(u) \\ \underline{K}(u, x, v, y) & \text{if } u \in U \text{ and } v \in V(u) \end{cases}$$

$$= \underline{K}(u, x, v, y).$$

Combining this with (2) and (3) yields $\underline{K} \leq N \leq \bar{K}$ everywhere. Hence Theorem 0.1(b) implies $N \in [K]$.

We conclude this section by characterizing the subdifferential of the partial conjugate.

THEOREM 4.3. The following conditions are equivalent:

- (a) $(u^*, x^*, v^*, y^*) \in \partial K(u, x, v, y)$
- (b) $(u, -y^*, v, -x^*) \in \partial W(u^*, y, v^*, x)$

PROOF. By (37.5) condition (b) is equivalent to

$$(u^*, y, v^*, x) \in \partial W^*(u, -y^*, v, -x^*).$$

But from the proof of Theorem 4.1 we know that $[W^*] = [H^* A^*]$ and $\text{range } A^* \cap \text{ri}(\text{dom } H^*) \neq \emptyset$. Hence by Theorem 1.3,

$$\partial W^*(u, -y^*, v, -x^*) = A \partial H^*(A^*(u, -y^*, v, -x^*)).$$

It follows that condition (b) is equivalent to the existence of points u^0 and v^0 such that

$$(v^0, u^*, y, u^0, v^*, x) \in \partial H^*(0, u, -y^*, 0, v, -x^*).$$

But by (37.5) and (37.4) this containment occurs if and only if $(v^0, u^*, y, u^0, y^*, u)$ is a saddle point of

$$H = \langle \cdot, (0, u, -y^*) \rangle - \langle \cdot, (0, v, -x^*) \rangle$$

and $H(v^0, u^*, y, u^0, v^*, x) \in R$. Therefore by the definition of H , condition (b) is equivalent to the existence of points u^0 and v^0 such that $k(u^0, x, v^0, y) \in R$ and

$$\begin{aligned} K(\bar{u}, \bar{x}, v^0, y) &= \langle \bar{u} - u, u^* \rangle - \langle v^0 - v, \bar{v}^* \rangle - \langle \bar{x} - x, x^* \rangle \\ &\leq K(u^0, x, v^0, y) - \langle u^0 - u, u^* \rangle - \langle v^0 - v, v^* \rangle \\ &\leq K(u^0, x, \bar{v}, \bar{y}) - \langle u^0 - u, \bar{u}^* \rangle - \langle \bar{v} - v, v^* \rangle - \langle \bar{y} - y, y^* \rangle \end{aligned}$$

for all $(\bar{v}, \bar{u}^*, \bar{y})$ and $(\bar{u}, \bar{v}^*, \bar{x})$. Now pick any $(v', y') \in \text{dom}_2 K$. Choosing $\bar{v} = v'$ and $\bar{y} = y'$ in the above condition implies

$$K(u^0, x, v', y') \geq \alpha + \langle u^0 - u, \bar{u}^* \rangle \text{ for all } \bar{u}^*,$$

where α is a certain real constant. Thus if u^0 were different from u , we would have $K(u^0, x, v', y') = +\infty$, contradicting $(v', y') \in \text{dom}_2 K$. Hence in the above condition we must have $u^0 = u$, and similarly $v^0 = v$. Therefore condition (b) is equivalent to $(K(u, x, v, y) \in R \text{ and})$

$$K(\dot{u}, x, v, \gamma) - \langle \dot{u} - u, u^\circ \rangle - \langle \dot{x} - x, x^\circ \rangle \leq K(u, x, v, \gamma), \quad \forall (\dot{u}, \dot{x})$$

and

$$K(u, x, v, \gamma) \leq K(u, x, \dot{v}, \dot{\gamma}) - \langle \dot{v} - v, v^\circ \rangle - \langle \dot{\gamma} - \gamma, \gamma^\circ \rangle, \quad \forall (\dot{v}, \dot{\gamma}).$$

But these conditions are clearly equivalent to (a).

§5. Generalized Saddle Programs

In this section we apply the results of §§1, 2 and 4 to the study of general concave-convex minimax problems. First we define the notions of a "generalized saddle program" and its "dual," and then we develop a whole perturbational duality theory for such programs. It may happen that the saddle functions defining these saddle programs are degenerate in the sense of being essentially purely convex or purely concave functions. In this event it can be shown (Example 5.1) that the present theory reduces essentially to Rockafellar's perturbational duality theory for generalized convex programs [48]. Recall, however, the general approach we are taking to minimax problems, namely that of dealing always with the whole equivalence class of saddle functions which give rise to a given minimax problem. Because of this, the proper definitions concerning generalized saddle programs involve many subtleties absent in the convex case, and the proofs in the accompanying perturbational duality theory are necessarily somewhat different and more complicated.

Ignoring technical details, we can outline the general approach as follows. Suppose we are given a minimax problem in the form of an equivalence class $[K_0]$ of saddle functions on $R^m \times R^n$. This minimax problem is first extended to a saddle program in the form of another equivalence class $[K]$ of saddle functions on $(R^p \times R^m) \times (R^q \times R^n)$, where the additional variables ranging over R^p and R^q correspond to "perturbations" of the original problem. The extension is such that $[K_0]$ is suitably "embedded" in $[K]$, i.e., the

saddle functions $(x, y) \rightarrow \tilde{K}(0, x, 0, y)$ for $\tilde{K} \in [K]$ are all required to belong to $[K_0]$. By a modification of the conjugacy correspondence, an equivalence class $[L]$ of saddle functions on $(R^m \times R^p) \times (R^n \times R^q)$ is then obtained from $[K]$. The saddle program given by $[L]$ is called the dual of the program given by $[K]$. Under a mild hypothesis on $[K]$, the saddle functions $(z, w) \rightarrow \tilde{L}(0, z, 0, w)$ for $\tilde{L} \in [L]$ all belong to a single equivalence class $[L_0]$. In this event the minimax problem given by $[L_0]$ is the dual of the minimax problem given by $[K_0]$. In this sense $[K_0]$ may have many such duals, since $[L]$ and hence $[L_0]$ depends not only on $[K_0]$ but also on the particular "perturbations" of $[K_0]$ introduced via $[K]$. This fact is one of the main features of the theory, since it allows one the flexibility of choosing perturbations which are appropriate for the purpose at hand (e.g. those for which the dual problem is manageable).

We proceed now with the formal development. A generalized saddle program $S(K)$ on $R^m \times R^n$ with perturbations in $R^p \times R^q$ is an equivalence class $[K]$ of closed proper saddle functions on $(R^p \times R^m) \times (R^q \times R^n)$. Each saddle function $\tilde{K}(0, \cdot, 0, \cdot)$ on $R^m \times R^n$, for \tilde{K} in $[K]$, is called an objective function of $S(K)$. The particular functions $\underline{K}(0, \cdot, 0, \cdot)$ and $\bar{K}(0, \cdot, 0, \cdot)$ are called the lower and upper objective functions, respectively. (Here, as usual, \underline{K} and \bar{K} denote the least and greatest elements of $[K]$, respectively.) A pair (x, y) is a feasible solution of $S(K)$ if and only if it is in the effective domain of every objective function of $S(K)$. It is not hard to show that this is equivalent to the condition that $(0, x, 0, y) \in \text{dom } K$. The optimal value in $S(K)$ exists

(and equals α) if and only if the saddle values of all the objective functions of $S(K)$ exist finitely and are equal (to α). A pair (x, y) is an optimal solution of $S(K)$ if and only if (x, y) is a saddle point of every objective function of $S(K)$ and $\underline{K}(0, x, 0, y) = \bar{K}(0, x, 0, y) \in \mathbb{R}$. It is not hard to show that if (x, y) is an optimal solution, then it is a feasible solution and the optimal value exists and equals $\tilde{K}(0, x, 0, y)$ for any \tilde{K} in $[K]$.

The program $S(K)$ is consistent (respectively strongly consistent) if and only if there exist points x and y such that $(0, x, 0, y) \in \text{dom } K$ (respectively $(0, x, 0, y) \in \text{ri}(\text{dom } K)$). Thus, $S(K)$ is consistent if and only if it has a feasible solution. Also, $S(K)$ is consistent whenever the optimal value in $S(K)$ exists.

We say that $S(K)$ has a well-defined primal problem if and only if all the objective functions belong to the same equivalence class, which we denote by $[K_0]$. In this event the primal problem of $S(K)$ is the minimax problem given by $[K_0]$ and hence the definitions of feasible solution, optimal value and optimal solution of $S(K)$ can be simplified, since equivalent saddle functions have the same effective domain, saddle value and saddle points. By Theorem 5.2 below, if $S(K)$ is strongly consistent then it has a well-defined primal problem which is in fact given by a closed proper equivalence class. More generally, for any (u, v) we say that the perturbation (u, v) in $S(K)$ is well-defined if and only if the saddle functions $\tilde{K}(u, \cdot, v, \cdot)$ on $\mathbb{R}^m \times \mathbb{R}^n$, for \tilde{K} in $[K]$, all belong to a single equivalence class, which we denote by $[K_{u,v}]$. Thus, $S(K)$ has a well-defined primal problem if and only if the perturbation $(0, 0)$ in $S(K)$ is well-defined (in which case $[K_{0,0}]$ is denoted simply by $[K_0]$).

Suppose $S(K)$ is a generalized saddle program on $R^m \times R^n$ with perturbations in $R^p \times R^q$, and let $[L]$ be the equivalence class of closed proper saddle functions obtained from $[K]$ via the relation

$$L(s, z, t, w) = -K^*(-z, s, -w, t) .$$

The generalized saddle program $S(L)$ on $R^p \times R^q$ with perturbations in $R^m \times R^n$ (strictly speaking, the equivalence class $[L]$) is the dual of $S(K)$. It is easy to show that the dual of $S(L)$ is $S(K)$. The program $S(K)$ has a well-defined dual problem if and only if the dual program $S(L)$ has a well-defined primal problem $[L_0]$, and in this event the dual problem of $S(K)$ is the primal problem of $S(L)$, i.e. the minimax problem given by $[L_0]$. Example 5.3 shows that a generalized saddle program can even be strongly consistent without having a well-defined dual problem. However, Lemma 5.4 furnishes conditions on $S(K)$ which ensure that the dual problem is well-defined.

For the remainder of §5 let $S(K)$ and $S(L)$ be a dual pair of generalized saddle programs, and for definiteness assume that K is concave-convex on $(R^p \times R^m) \times (R^q \times R^n)$. Thus, L is convex-concave on $(R^m \times R^p) \times (R^n \times R^q)$. Also, let concave-convex functions P_1 and P_2 be defined on $R^p \times R^q$ by

$$P_1(u, v) = \sup_x \inf_y \underline{K}(u, x, v, y)$$

and

$$P_2(u, v) = \inf_y \sup_x \bar{K}(u, x, v, y) ,$$

and let convex-concave functions Q_1 and Q_2 be defined on $R^m \times R^n$ by

$$Q_1(s, t) = \sup_w \inf_z \underline{L}(s, z, t, w)$$

and

$$Q_2(s, t) = \inf_z \sup_w \bar{L}(s, z, t, w) .$$

Finally, let linear transformations $A_1: R^p \times R^m \rightarrow R^p$, $A_2: R^q \times R^n \rightarrow R^q$, $B_1: R^m \times R^p \rightarrow R^m$ and $B_2: R^n \times R^q \rightarrow R^n$ be defined by

$$A_1(u, x) = u, \quad B_1(s, z) = s ,$$

$$A_2(v, y) = v, \quad B_2(t, w) = t ,$$

and put $A = A_1 \times A_2$ and $B = B_1 \times B_2$. Observe that $A^* = A_1^* \times A_2^*$ and $B^* = B_1^* \times B_2^*$, where

$$A_1^*(z) = (z, 0), \quad B_1^*(x) = (x, 0)$$

$$A_2^*(w) = (w, 0), \quad B_2^*(y) = (y, 0) .$$

The saddle functions P_1 and P_2 are called the lower and upper perturbation functions of $S(K)$, respectively. A pair (z, w) is a Kuhn-Tucker vector for $S(K)$ if and only if

$$P_1(0, 0) = P_2(0, 0) = \alpha \in R$$

and

$$\langle u, z \rangle + P_2(u, 0) \leq \alpha \leq P_1(0, v) + \langle v, w \rangle$$

for each (u, v) . Observe that $P_1(0, 0) = P_2(0, 0) = \alpha \in R$ if and only if the optimal value in $S(K)$ exists and equals α . It is not hard to show using (37.4.1) that if P_1 and P_2 belong to a proper equivalence class $[P]$, then (z, w) is a Kuhn-Tucker vector for $S(K)$ if and only if $-(z, w) \in \partial P(0, 0)$. Kuhn-Tucker vectors for $S(L)$ are defined similarly by using the lower and upper perturbation functions of $S(L)$, i.e., Q_1 and Q_2 . These Kuhn-Tucker vectors can be interpreted as generalized "equilibrium price vectors" for a certain type of two-stage, two-person zero-sum game in much the same way as in [48, pp. 295-296].

The following example shows that the foregoing framework of dual pairs of generalized saddle programs includes as a special case Rockafellar's dual pairs of generalized convex programs.

EXAMPLE 5.1. Let $F: R^q \rightarrow R^n$ be a closed proper convex bifunction, and let (P) and (P^*) denote the generalized convex program and its dual which correspond to F and its adjoint bifunction $F^*: R^n \rightarrow R^q$. Define $K(u, x, v, y) = (Fv)(y)$ for every $(u, x) \in R^p \times R^m$ and $(v, y) \in R^q \times R^n$ (here p and m can be any positive integers). Then K is a closed proper concave-convex function. It can be verified as an instructive exercise that the concepts defined above for the program $S(K)$ and its dual $S(L)$ exactly "mirror" the like-named concepts from Rockafellar [48] for (P) and (P^*) . For example, $S(K)$ (resp. $S(L)$) is consistent or strongly consistent according as (P) (resp. (P^*)) is consistent or strongly consistent; and so on. Furthermore, it can be seen that the Lagrangian associated with $S(K)$ and $S(L)$, which will be introduced following Theorem 5.6, exactly mirrors the Lagrangian associated with (P) and (P^*) . The fact that all the various concepts associated with (P) and (P^*) are reflected in this program $S(K)$ and its dual furnishes a general means of converting examples from convex programming into examples in saddle programming which exhibit the corresponding pathological behavior.

The following theorem gives a simple condition under which the perturbation (u, v) in $S(K)$ is well-defined, and Corollary 5.2.1 is the main existence result concerning optimal solutions of $S(K)$.

THEOREM 5.2. Assume there exist points x and y such that $(u, x, v, y) \in \text{ri}(\text{dom } K)$. Then the perturbation (u, v) in $S(K)$ is well-defined. In fact, the equivalence class $[K_{u,v}]$ is closed and proper with least and greatest members $\underline{K}(u, \cdot, v, \cdot)$ and $\bar{K}(u, \cdot, v, \cdot)$ respectively, and

$$\text{ri}(\text{dom } K_{u,v}) = \{(x, y) \mid (u, x, v, y) \in \text{ri}(\text{dom } K)\}$$

where "ri" can be deleted or replaced by "cl" throughout the identity.

PROOF. Define linear transformations $T_1: R^m \rightarrow R^p \times R^m$ and $T_2: R^n \rightarrow R^q \times R^n$ by $T_1(x) = (0, x)$ and $T_2(y) = (0, y)$, and put $T = T_1 \times T_2$. Define a function H by

$$H(u', x', v', y') = K(u' + u, x', v' + v, y') .$$

Clearly, H is closed proper concave-convex and $\text{dom } H = \text{dom } K - \{(u, 0, v, 0)\}$.

Thus the hypothesis on (u, v) is equivalent to $\text{range } T \cap \text{ri}(\text{dom } H) \neq \emptyset$, and

hence Theorem 1.2 implies various facts about the equivalence class $[HT]$.

Since $HT = K(u, \cdot, v, \cdot)$, these facts convert easily into the assertions about

$[K_{u,v}]$. The formulas for $\text{ri}(\text{dom } K_{u,v})$ and $\text{cl}(\text{dom } K_{u,v})$ follow by (6.7).

COROLLARY 5.2.1. Assume $S(K)$ is strongly consistent. Then any of the following three conditions implies that the set of optimal solutions of $S(K)$ is a product of nonempty closed convex sets:

- (a) $\text{rec cone}_j K$ is a subspace for $j = 1$ and 2 ;
- (b) $\text{dom } K_0$ is bounded;
- (c) There is a pair $(\bar{x}, \bar{y}) \in \text{dom } K_0$ such that for each $\alpha \in R$ the level sets $\{x \mid K_0(x, \bar{y}) \geq \alpha\}$ and $\{y \mid K_0(\bar{x}, y) \leq \alpha\}$ are bounded.

PROOF. By Theorem 5.2, $S(K)$ has a well-defined primal problem and $[K_0]$ is closed and proper. By Lemmas 0.5 and 0.6, $(0, 0) \in \text{ri}(\text{dom}(K_0)^*)$ whenever (a), (b) or (c) holds. The conclusion now follows from (37.5.3).

Before proceeding any further, it might be well to illustrate some of the pathology which is possible in a dual pair of generalized saddle programs. The next example exhibits a program $S(K)$ having the following properties:

(i) every perturbation in $S(K)$ is well-defined (so a fortiori $S(K)$ has a well-defined primal problem); (ii) the lower and upper perturbation functions of $S(K)$ fail to be equivalent; (iii) the dual program is consistent; and (iv) $S(K)$ fails to have a well-defined dual problem.

EXAMPLE 5.3. Suppose J is a closed proper concave-convex function on $R^m \times R^n$. Put $p = m$ and $q = n$ and define $K(u, x, v, y) = J(x - u, y - v)$. Let T_1 and T_2 be linear transformations given by $T_1(u, x) = x - u$ and $T_2(v, y) = y - v$, and put $T = T_1 \times T_2$. Since $\text{range } T \cap \text{ri}(\text{dom } J) \neq \emptyset$ trivially, Theorem 1.2 implies that $K = JT$ is closed and proper with $\text{ri}(\text{dom } K) = T^{-1} \text{ri}(\text{dom } J)$. By Theorem 5.2 it follows that for each $(u, v) \in R^p \times R^q$ the perturbation (u, v) in the program $S(K)$ is well-defined. It is easy to compute that $P_1(u, v) = \sup \inf J = \bar{J}^*(0, 0)$ and $P_2(u, v) = \inf \sup \bar{J} = -\underline{J}^*(0, 0)$. Hence $P_1 \sim P_2$ if and only if $\underline{J}^*(0, 0) = \bar{J}^*(0, 0)$. Now suppose J is such that $\text{dom } J^*$ is bounded. Then Lemma 2.6 and Theorem 2.5 imply that $[K^*] = [T^* J^*]$, $\text{dom } K^* = T^* \text{dom } J^*$, and (since $T_1^*(s) = (-s, s)$, $T_2^*(t) = (-t, t)$) the least and greatest members of $[K^*]$ are

$$\bar{K}^*(z, s, w, t) = \sup_{\{s \mid -s=z\}} \inf_{\{t \mid -t=w\}} \bar{J}^*(s, t)$$

and

$$\bar{K}^*(z, s, w, t) = \inf_{\{t \mid -t=w\}} \sup_{\{s \mid -s=z\}} \bar{J}^*(s, t) .$$

Since $\underline{L}(s, z, t, w) = -\bar{K}^*(-z, s, -w, t)$ and $\bar{L}(s, z, t, w) = -\underline{K}^*(-z, s, -w, t)$,
this implies that

$$\underline{L}(s, z, t, w) = \begin{cases} -\bar{J}^*(s, t) & \text{if } s = z \text{ and } t = w \\ +\infty & \text{if } s \neq z \text{ and } t = w \\ -\infty & \text{if } t \neq w \end{cases}$$

and

$$\bar{L}(s, z, t, w) = \begin{cases} -\underline{J}^*(s, t) & \text{if } s = z \text{ and } t = w \\ -\infty & \text{if } s = z \text{ and } t \neq w \\ +\infty & \text{if } s \neq z . \end{cases}$$

From these formulas it follows that, for each $(s, t) \in \text{dom } J^*$, the perturbation (s, t) in $S(L)$ is well-defined if and only if $\underline{J}^*(s, t) = \bar{J}^*(s, t)$. In view of all these facts, in order to obtain properties (i) through (iv) we need only specify a J such that $\text{dom } J^*$ is bounded, $(0, 0) \in \text{dom } J^*$, and $\underline{J}^*(0, 0) \neq \bar{J}^*(0, 0)$. But for this it suffices to take $[J]$ to be the conjugate of the equivalence class used in Examples 1.10 and 1.11.

By Theorem 5.2, $S(K)$ has a well-defined dual problem $[L_0]$ whenever $S(L)$ is strongly consistent. The next lemma dualizes this useful condition.

LEMMA 5.4. The program $S(L)$ is strongly consistent if and only if

$$(\text{rec}_1 K)(0, x) \geq 0 \text{ implies } (\text{rec}_1 K)(0, -x) \geq 0$$

and

$$(\text{rec}_2 K)(0, y) \leq 0 \text{ implies } (\text{rec}_2 K)(0, -y) \leq 0 .$$

PROOF. Observe that $\underline{L}(0, z, 0, w) = -\bar{K}^* A^* (-z, -w)$. Hence $S(L)$ is strongly consistent if and only if $\text{range } A^* \cap \text{ri}(\text{dom } K^*) \neq \emptyset$. Now apply the equivalence between (a) and (c) of Lemma 1.9.

The following theorem and its corollaries show that, when the dual program is strongly consistent, much information about the dual program may be converted into information about the primal program.

THEOREM 5.5. Assume $S(L)$ is strongly consistent. Then P_1 and P_2 belong to the closed proper equivalence class $[P] = [-(L_0)^*]$ and $\text{dom } P \subset \text{Adom } K$.

PROOF. By Theorem 5.2, $\underline{L}(0, \cdot, 0, \cdot)$ is the least member of $[L_0]$, which is closed and proper. Hence $-\underline{L}(0, -z, 0, -w) = -\underline{L}_0(-z, -w) = \underline{-(L_0)^*}^*(z, w)$. But as noted in the proof of Lemma 5.4, $S(L)$ is strongly consistent if and only if $\text{range } A^* \cap \text{ri}(\text{dom } K^*) \neq \emptyset$, and $\bar{K}^* A^*(z, w) = -\underline{L}(0, -z, 0, -w)$. Hence Theorem 1.8 implies that the equivalence class $[AK]$ is well-defined and equals $\underline{-(L_0)^*}$, and $\text{dom } AK \subset \text{Adom } K$. Now observe that

$$P_1(u, v) = \sup_{A_1^{-1}\{u\}} \inf_{A_2^{-1}\{v\}} K, \quad P_2(u, v) = \inf_{A_2^{-1}\{v\}} \sup_{A_1^{-1}\{u\}} \bar{K} .$$

Thus P_1 and P_2 belong to $[AK]$. Taking $[P] = [AK]$, the theorem follows.

COROLLARY 5.5.1. Assume $S(L)$ is strongly consistent, and let $[P]$ be the equivalence class containing P_1 and P_2 . Then the following conditions on $(z, w) \in R^p \times R^q$ are equivalent:

- (i) (z, w) is an optimal solution of $S(L)$;
- (ii) (z, w) is a Kuhn-Tucker vector for $S(K)$;

(iii) $-(z, w) \in \partial P(0, 0)$;

(iv) $(-z, 0, -w, 0) \in \partial K(0, x, 0, y)$ for some $(x, y) \in R^m \times R^n$.

PROOF. By (37.5) and Theorem 5.5, (i) is equivalent to $(z, w) \in \partial(-P)(0, 0)$, which is equivalent to (iii) . Since P_1 and P_2 belong to $[P]$, (37.4.1) implies that $\partial P(0, 0) = \partial_1 P_2(0, 0) \times \partial_2 P_1(0, 0)$ and $P_1(0, 0) = P_2(0, 0) = \alpha$. Also, (37.4) implies $\text{dom } \partial P \subset \text{dom } P$, so that α is finite. From these facts it follows that (iii) is equivalent to (ii) . Finally, observe that (37.5) implies (iii) is equivalent to $(0, 0) \in \partial P^*(-z, -w)$. Since $[P^*] = [K^* A^*]$ by the proof of Theorem 5.5, Theorem 1.3 implies that $\partial P^*(-z, -w) = A \partial K^*(A^*(-z, -w))$. Hence $(0, 0) \in \partial P^*(-z, -w)$ is equivalent to the existence of $(u, x, v, y) \in \partial K^*(-z, 0, -w, 0)$ such that $A_1(u, x) = 0$ and $A_2(v, y) = 0$. By the definitions of A_1 and A_2 and (37.5), this last condition is equivalent to (iv) .

COROLLARY 5.5.2. Assume $S(L)$ is strongly consistent, and let $[P]$ be the equivalence class containing P_1 and P_2 . Then

$$\sup \inf L_0 = \underline{P}(0, 0) \leq \overline{P}(0, 0) = \inf \sup L_0 .$$

Furthermore, for any $\tilde{P} \in [P]$,

$$\sup \inf L_0 = \lim_{v \rightarrow 0} \inf \tilde{P}(0, v)$$

except when the left hand side is $-\infty$ and the right hand side is $+\infty$, and similarly

$$\lim_{u \rightarrow 0} \sup \tilde{P}(u, 0) = \inf \sup L_0$$

except when the left hand side is $-\infty$ and the right hand side is $+\infty$.

PROOF. Clearly $\sup \inf L_0 = -(\overline{L_0})^*(0,0)$, which equals $\underline{P}(0,0)$ by Theorem 5.5. Now let $\tilde{P} \in [P]$ be given. By Theorem 0.1, $\underline{P}(0,0) = (\text{cl}_2 \tilde{P})(0,0)$, which by definition equals $(\text{cl } \tilde{P}(0, \cdot))(0)$. Observe that in general, for a convex function f one has $(\text{cl } f)(x) = \lim_{y \rightarrow x} \inf f(y)$ except when the left hand side is $-\infty$ and the right hand side is $+\infty$. Now apply this to the case at hand. A concave analogue of this argument yields the other assertion.

COROLLARY 5.5.3. Assume $S(L)$ is strongly consistent. If the optimal value in $S(L)$ exists and equals α , then the optimal value in $S(K)$ exists and equals α .

PROOF. Since the saddle functions $\tilde{L}(0, \cdot, 0, \cdot)$ for \tilde{L} in $[L]$ are all equivalent to L_0 , $\sup \inf L_0 = Q_1(0,0)$ and $\inf \sup L_0 = Q_2(0,0)$. Also, the optimal value in $S(L)$ exists and equals α if and only if $Q_1(0,0) = Q_2(0,0) = \alpha \in \mathbb{R}$. Since $\underline{P} \leq P_1 \leq P_2 \leq \bar{P}$, the assertion now follows from Corollary 5.5.2.

For stating duality results it would be very nice if the domain inclusion in Theorem 5.5 could be strengthened to $\text{Ari}(\text{dom } K) \subset \text{dom } P \subset \text{Adom } K$. However, Example 5.14 shows that this does not hold without some stronger hypothesis. The next corollary characterizes when the additional inclusion holds.

COROLLARY 5.5.4. Assume $S(L)$ is strongly consistent, and let $[P]$ be the equivalence class containing P_1 and P_2 . In order that

$$\text{Ari}(\text{dom } K) \subset \text{dom } P \subset \text{Adom } K,$$

it is necessary and sufficient that

$$(\text{rec}_1 L)(0, z) = (\text{rec}_1 L_0)(z), \quad \forall z \in \mathbb{R}^p$$

and

$$(\text{rec}_2 L)(0, w) = (\text{rec}_2 L_0)(w), \quad \forall w \in R^q.$$

PROOF. As in the proof of Theorem 5.5, $S(L)$ strongly consistent is equivalent to $\text{range } A^* \cap \text{ri}(\text{dom } K^*) \neq \emptyset$ and hence from Theorem 1.8 we have $[P] = [AK]$. By Lemma 1.5 (with the help of (6.3.1)) it follows that for $j = 1$ and 2 , $\text{ri}(A_j \text{ dom}_j K) \subset \text{dom}_j P \subset A_j \text{ dom}_j K$ if and only if $\text{rec}_j(K^* A^*) = (\text{rec}_j K^*) A_j^*$. But the identities $-(\text{rec}_1 L)(0, -z) = (\text{rec}_1 K^*)(A_1^* z)$ and $-(\text{rec}_1 L_0)(-z) = (\text{rec}_1 K^* A^*)(z)$ can be verified, along with similar identities for $j = 2$. The corollary then follows.

The next theorem brings the results of §2 to bear on the perturbations in $S(K)$. It gives conditions under which any "perturbed primal problem" (represented by $[K_{u,v}]$) has a "solution," provided only that the perturbation (u, v) lies in $\text{Ari}(\text{dom } K)$. The boundedness assumption on $\text{dom } K_0$ can be relaxed; we have used this hypothesis for simplicity.

THEOREM 5.6. Assume $S(K)$ is strongly consistent and $\text{dom } K_0$ is bounded. Then $S(L)$ is strongly consistent, P_1 and P_2 are respectively the least and greatest members of a closed proper equivalence class $[P]$, and $\text{dom } P = \text{Adom } K$. Moreover, for each $(u, v) \in \text{Ari}(\text{dom } K)$ the "perturbed objective function" $K_{u,v}$ has a nonempty bounded set of saddle points, and each such saddle point (x, y) yields attainment of the minimax extrema in $P_1(u, v)$ and $P_2(u, v)$ and satisfies $\tilde{P}(u, v) = \tilde{K}(u, x, v, y)$, $\forall \tilde{P} \in [P], \forall \tilde{K} \in [K]$.

PROOF. By strong consistency there exist \bar{x} and \bar{y} such that $(0, \bar{x}, 0, \bar{y}) \in \text{ri}(\text{dom } K)$. Observe that $\text{rec } K_0(\bar{x}, \cdot)(y) \leq (\text{rec}_2 K_0)(y) \leq (\text{rec}_2 K)(0, y)$

for any y . Since $\text{dom } K_0(\bar{x}, \cdot) = \text{dom}_2 K_0$ is bounded by hypothesis, it follows that $y = 0$ is the only solution of $(\text{rec}_2 K)(0, y) \leq 0$. Similarly, $x = 0$ is the only solution of $(\text{rec}_1 K)(0, x) \geq 0$. Therefore Lemma 5.4 implies $S(L)$ is strongly consistent. Now observe that our hypotheses allow us to refine the proof of Theorem 5.5. Specifically, Lemma 2.6 and Theorem 5.2 can be used to verify that the hypotheses of Theorem 2.5 (and hence Theorem 2.4) are satisfied. The fact that A_1 and A_2 are onto implies by Theorem 2.5 that the least and greatest members of $[AK]$ are actually P_1 and P_2 . The second assertion follows from Theorem 2.4.

Some duality theorems for the programs $S(K)$ and $S(L)$ can be deduced from the results already given. However, much more can be derived by introducing a "Lagrangian" saddle function and studying its relationships with $S(K)$ and $S(L)$. Consequently we shall now proceed with this.

Define the function M on $(R^m \times R^q) \times (R^n \times R^p)$ by

$$M(x, w, y, z) = \sup_u \inf_v \{ \langle u, z \rangle + \langle v, w \rangle + K(u, x, v, y) \} .$$

Then $M(x, w, y, z) = -W(-z, y, -w, x)$, where W belongs to the partial conjugate of $[K]$ in u and v . Hence it follows from Theorem 4.1 that M is closed proper concave-convex and depends only on $[K]$, and that M is polyhedral whenever K is polyhedral. The equivalence class containing M is called the Lagrangian of $S(K)$. Similarly, the Lagrangian of $S(L)$ is the equivalence class containing the function N on $(R^p \times R^n) \times (R^q \times R^m)$ given by

$$N(z, y, w, x) = \sup_t \inf_s \{ \langle s, x \rangle + \langle t, y \rangle + L(s, z, t, w) \} .$$

Note that $N(z, y, w, x) = -Q(-x, w, -y, z)$, where Q belongs to the partial conjugate of $[L]$ in s and t . The following theorem reveals the connection between M and N .

THEOREM 5.7. The saddle functions $(x, w, y, z) \rightarrow M(x, w, y, z)$ and $(x, w, y, z) \rightarrow N(z, y, w, x)$ are equivalent.

PROOF. Let $P(x, w, y, z) = N(z, y, w, x)$. We must show $P \sim M$, or what is the same thing, $[P] = [M]$. Since $[L]$ is obtained from $[K^*]$ via the relation $L(s, z, t, w) = -K^*(-z, s, -w, t)$, it follows by (36.1) and Theorem 0.1(b) that $[L]$ contains the function \tilde{L} given by

$$\begin{aligned}\tilde{L}(s, z, t, w) &= -\sup_y \inf_x \inf_u \sup_v \{ \langle u, -z \rangle + \langle x, s \rangle + \langle v, -w \rangle + \langle y, t \rangle - M(x, w, y, z) \} \\ &= -\sup_y \inf_x \{ \langle x, s \rangle + \langle y, t \rangle - M(x, w, y, z) \} .\end{aligned}$$

Letting $[H]$ denote the partial conjugate of $[M]$ in x and y , this means $[L] = [-H]$. Now by Theorem 4.1 the equivalence class $[N]$ depends only on $[L]$. Hence $[L] = [-H]$ implies that $[N]$ contains the function \tilde{N} given by

$$\tilde{N}(z, y, w, x) = \sup_t \inf_s \{ \langle s, x \rangle + \langle t, y \rangle - \tilde{H}(s, z, t, w) \} ,$$

where \tilde{H} is any element of $[H]$. Therefore $[P]$ contains the function \tilde{P} given by $\tilde{P}(x, w, y, z) = \tilde{N}(z, y, w, x)$. But \tilde{P} belongs to the partial conjugate of $[H]$ in s and t , which by Theorem 4.2 is the same as $[M]$. This shows that $[P] = [M]$.

The content of this theorem is that the Lagrangian of $S(L)$ is just the Lagrangian of $S(K)$ with the variables reversed. So henceforth we shall deal

only with $[M]$ and refer to it as the Lagrangian of the dual pair or simply the Lagrangian. By the results of §4, the partial conjugacy correspondence is one-to-one and symmetric among closed proper equivalence classes. From this it follows that a dual pair of generalized saddle programs completely determines its Lagrangian and vice versa.

THEOREM 5.8. Let $[M]$ denote the Lagrangian. Then $[-K]$ is the partial conjugate of $[M]$ in w and z , and $[-L]$ is the partial conjugate of $[M]$ in x and y .

PROOF. The second assertion follows from the fact (indicated at the beginning of the proof of Theorem 5.7) that $[L]$ contains the function \tilde{L} given by

$$\tilde{L}(s, z, t, w) = -\sup_y \inf_x \{ \langle x, s \rangle + \langle y, t \rangle - M(x, w, y, z) \} .$$

To see the first assertion, notice that $M(x, w, y, z) = -W(-z, y, -w, x)$ together with Theorem 4.2 imply that $[K]$ contains the function \tilde{K} given by

$$\begin{aligned} \tilde{K}(u, x, v, y) &= \sup_w \inf_z \{ \langle z, u \rangle + \langle w, v \rangle - W(z, y, w, x) \} \\ &= -\inf_w \sup_z \{ \langle z, u \rangle + \langle w, v \rangle - M(x, w, y, z) \} . \end{aligned}$$

This theorem yields representations of the primal and dual objective functions in terms of the Lagrangian. Recall from Theorem 5.2 that in the presence of strong consistency all the objective functions are closed, proper and equivalent.

COROLLARY 5.8.1. Let \tilde{M} be any member of the Lagrangian. Then both of the functions

$$(x, y) \rightarrow \sup_w \inf_z \tilde{M}(x, w, y, z) \text{ and } (x, y) \rightarrow \inf_z \sup_w \tilde{M}(x, w, y, z)$$

are objective functions of $S(K)$, and both of the functions

$$(z, w) \rightarrow \sup_x \inf_y \tilde{M}(x, w, y, z) \text{ and } (z, w) \rightarrow \inf_y \sup_x \tilde{M}(x, w, y, z)$$

are objective functions of $S(L)$.

PROOF. Since $[-K]$ is the partial conjugate of $[M]$ in w and z , it follows that $[K]$ contains the functions K_1 and K_2 given by

$$K_1(u, x, v, y) = -\inf_w \sup_z \{ \langle z, u \rangle + \langle w, v \rangle - \tilde{M}(x, w, y, z) \} ,$$

$$K_2(u, x, v, y) = -\sup_z \inf_w \{ \langle z, u \rangle + \langle w, v \rangle - \tilde{M}(x, w, y, z) \} .$$

Fixing $(u, v) = (0, 0)$ in these functions yields the indicated objective functions of $S(K)$. The other assertion is proved similarly.

The next theorem reveals more of the close relationship between $[K]$, $[L]$ and $[M]$.

THEOREM 5.9. The equivalence class $[M^*]$ conjugate to the Lagrangian coincides with both the partial conjugate of $[K]$ in x and y and also, except for reversing the orders of the variables, the partial conjugate of $[L]$ in z and w .

PROOF. The partial conjugate of $[K]$ in x and y contains the function H given by

$$H(s, v, t, u) = \sup_y \inf_x \{ \langle x, s \rangle + \langle y, t \rangle - \tilde{K}(u, x, v, y) \} .$$

Here \tilde{K} can be any member of $[K]$. By choosing the particular \tilde{K} appearing in the proof of Theorem 5.8 we obtain

$$H(s, v, t, u) = \sup_y \inf_x \inf_w \sup_z \{ \langle x, s \rangle + \langle y, t \rangle + \langle z, u \rangle + \langle w, v \rangle - M(x, w, y, z) \} .$$

But (by (36.1) and Theorem 0.1(b)) this function belongs to $[M^*]$. On the other hand, the partial conjugate of $[L]$ in z and w contains the function J given by

$$J(u, t, v, s) = \sup_z \inf_w \{ \langle z, u \rangle + \langle w, v \rangle - \tilde{L}(s, z, t, w) \} ,$$

where we can take the \tilde{L} to be the one in the proof of Theorem 5.8. Thus

$$J(u, t, v, s) = \sup_z \inf_w \sup_y \inf_x \{ \langle z, u \rangle + \langle w, v \rangle + \langle x, s \rangle + \langle y, t \rangle - M(x, w, y, z) \} .$$

But this formula for J shows that the function $(s, v, t, u) \rightarrow J(u, t, v, s)$ belongs to $[M^*]$.

This theorem allows us to obtain information about the perturbation functions and optimal values by means of the conjugate of the Lagrangian.

COROLLARY 5.9.1. If there exist points u and v such that $(0, v, 0, u) \in \text{ri}(\text{dom } M^*)$, then $-P_1$ and $-P_2$ belong to a closed proper equivalence class which contains the upper and lower conjugate of every objective function of $S(L)$. Similarly, if there exist points s and t such that $(s, 0, t, 0) \in \text{ri}(\text{dom } M^*)$, then $-Q_1$ and $-Q_2$ belong to a closed proper equivalence class which contains the upper and lower conjugate of every objective function of $S(K)$.

PROOF. Assume $(0, v, 0, u) \in \text{ri}(\text{dom } M^*)$ for some u and v . Then Theorem 5.2 implies that the functions $(v, u) \rightarrow \tilde{M}^*(0, v, 0, u)$ for $\tilde{M}^* \in [M^*]$ all belong to a single closed proper equivalence class. By Theorem 5.9 this implies that the functions

$$(v, u) \rightarrow -P_1(u, v) = \inf_x \sup_y \{ \langle x, 0 \rangle + \langle y, 0 \rangle - \underline{K}(u, x, v, y) \} ,$$

$$(v, u) \rightarrow -P_2(u, v) = \sup_y \inf_x \{ \langle x, 0 \rangle + \langle y, 0 \rangle - \bar{K}(u, x, v, y) \} ,$$

$$(v, u) \rightarrow \overline{(\tilde{L}(0, \cdot, 0, \cdot))^*}(u, v) = \inf_w \sup_z \{ \langle z, u \rangle + \langle w, v \rangle - \tilde{L}(0, z, 0, w) \} ,$$

$$(v, u) \rightarrow \underline{(\tilde{L}(0, \cdot, 0, \cdot))^*}(u, v) = \sup_z \inf_w \{ \langle z, u \rangle + \langle w, v \rangle - \tilde{L}(0, z, 0, w) \}$$

are equivalent, closed and proper. The second assertion follows similarly.

COROLLARY 5.9.2. If the saddle value of the Lagrangian exists and equals α , where $\alpha \in \mathbb{R}$, then the optimal values in $S(K)$ and $S(L)$ exist and equal α .

PROOF. The saddle value of M exists and equals α if and only if $\underline{M}^*(0, 0, 0, 0) = \bar{M}^*(0, 0, 0, 0) = -\alpha$. By Theorem 0.1(b) this is equivalent to $\tilde{M}^*(0, 0, 0, 0) = -\alpha$ for every $\tilde{M}^* \in [M^*]$. For $i = 1$ and 2 , $-P_i(u, v) = H_i(0, v, 0, u)$ for a certain member H_i of the partial conjugate of $[K]$ in x and y (cf. proof of Corollary 5.9.1), and similarly $-Q_i(s, t) = J_i(0, t, 0, s)$ for a certain member J_i of the partial conjugate of $[L]$ in z and w . Hence Theorem 5.9 implies that $-P_i(0, 0) = -Q_i(0, 0) = -\alpha$ for $i = 1, 2$.

For most of our remaining results we shall need the notion of a "stable" optimal solution (cf. [46], [47]). For each $x \in \mathbb{R}^m$ define the function f_x on

R^q by $f_x(v) = \inf K(0, x, v, \cdot)$, and for each $y \in R^n$ define the function g_y on R^p by $g_y(u) = \sup \bar{K}(u, \cdot, 0, y)$. It follows easily from (5.7) that f_x is convex and g_y is concave. An optimal solution (x, y) of $S(K)$ is said to be stable if and only if the directional derivative function

$$v \rightarrow f'_x(0;v) = \lim_{\lambda \downarrow 0} \frac{f_x(\lambda v) - f_x(0)}{\lambda}$$

is never $-\infty$ and the directional derivative function

$$u \rightarrow g'_y(0;u) = \lim_{\lambda \downarrow 0} \frac{g_y(\lambda u) - g_y(0)}{\lambda}$$

is never $+\infty$. As noted below in Lemma 5.10, (x, y) is an optimal solution of $S(K)$ if and only if $f_x(0) = g_y(0) \in R$. Hence by (23.1) the directional derivatives mentioned in this definition exist ($+\infty$ and $-\infty$ being allowed as limits). Stable optimal solutions of $S(L)$ are defined similarly, using the functions $h_w(s) = \inf \underline{L}(s, \cdot, 0, w)$ and $k_z(t) = \sup \bar{L}(0, z, t, \cdot)$.

LEMMA 5.10. Let $(x, y) \in R^m \times R^n$. Then (x, y) is an optimal solution of $S(K)$ if and only if $f_x(0) = g_y(0) \in R$, and in this event (x, y) is stable if and only if f_x and g_y are subdifferentiable at the origin.

PROOF. The first equivalence follows from the definitions and the fact that $\underline{K} \leq \tilde{K} \leq \bar{K}$ for each $\tilde{K} \in [K]$. The second equivalence then follows by (23.2) and (23.3).

We refer to the six equivalent conditions in the next theorem as the extremality conditions associated with the dual pair of programs $S(K)$ and $S(L)$.

THEOREM 5.11. For $(x, y) \in R^m \times R^n$ and $(z, w) \in R^p \times R^q$, each of the following conditions are equivalent:

- (a) $(-z, 0, -w, 0) \in \partial K(0, x, 0, y)$;
- (b) $(-x, 0, -y, 0) \in \partial L(0, z, 0, w)$;
- (c) $(0, 0, 0, 0) \in \partial M(x, w, y, z)$;
- (d) (x, w, y, z) is a saddle point of the Lagrangian;
- (e) $-(z, w) \in \partial g_y(0) \times \partial f_x(0)$ and $f_x(0) = g_y(0) \in R$;
- (f) $-(x, y) \in \partial h_w(0) \times \partial k_z(0)$ and $h_w(0) = k_z(0) \in R$.

PROOF. Observe that $-M(x, -w, y, -z) = W(z, y, w, x)$, where W belongs to the partial conjugate of $[K]$ in u and v . Also, Theorem 4.3 implies (a) is equivalent to $(0, 0, 0, 0) \in \partial W(-z, y, -w, x)$. By (37.4) it follows that (a) is equivalent to (c). Trivially, (c) is equivalent to (d). Observe from (37.5) that (a) is equivalent to $(0, x, 0, y) \in \partial K^*(-z, 0, -w, 0)$. By (37.4) and the relation $L(s, z, t, w) = -K^*(-z, s, -w, t)$ it follows that this last condition is equivalent to (b). By (37.4) and (37.4.1), (a) occurs if and only if $\bar{K}(0, x, 0, y) = \underline{K}(0, x, 0, y) = \alpha \in R$ and

$$\langle u, z \rangle + \bar{K}(u, \bar{x}, 0, y) \leq \alpha \leq \underline{K}(0, x, v, \bar{y}) + \langle v, w \rangle, \quad \forall u \quad \forall \bar{x} \quad \forall v \quad \forall \bar{y} .$$

But this occurs if and only if $g_y(0) = f_x(0) = \alpha \in R$ and

$$\langle u, z \rangle + g_y(u) \leq \alpha \leq f_x(v) + \langle v, w \rangle, \quad \forall u \quad \forall v .$$

These last conditions are equivalent to (e). Similarly, (b) is equivalent to (f).

An especially sharp duality theorem now follows easily.

THEOREM 5.12. The program $S(K)$ has a stable optimal solution if and only if $S(L)$ does, and in this case the two optimal values are equal.

PROOF. The equivalence follows from Lemma 5.10 and its dualized version, together with conditions (e) and (f) of Theorem 5.11. The optimal values are equal by Corollary 5.9.2 and condition (d) of Theorem 5.11.

From this theorem we can obtain a criterion for the nonexistence of stable optimal solutions.

COROLLARY 5.12.1. Assume that the optimal value in $S(K)$ exists and equals α . If either $\lim_{\lambda \downarrow 0} \frac{1}{\lambda} \{P_1(0, \lambda v) - \alpha\} = -\infty$ for some v or $\lim_{\lambda \downarrow 0} \frac{1}{\lambda} \{P_2(\lambda u, 0) - \alpha\} = +\infty$ for some u , then neither $S(K)$ nor $S(L)$ has a stable optimal solution.

PROOF. Suppose (x, y) is an optimal solution of $S(K)$. Then $f_x(0) = g_y(0) \in \mathbb{R}$. Notice that

$$f_x \leq \sup_{\bar{x}} f_{\bar{x}} = P_1(0, \cdot) \text{ and } P_2(\cdot, 0) = \inf_{\bar{y}} g_{\bar{y}} \leq g_y.$$

Hence the hypothesis implies either $f'_x(0; v) = -\infty$ for some v or $g'_y(0; u) = +\infty$ for some u . This means that (x, y) is not stable. Thus $S(K)$ has no stable optimal solution, and by the theorem neither does $S(L)$.

The next theorem relates stable optimal solutions to the extremality conditions.

THEOREM 5.13. The pair (x, y) is a stable optimal solution of $S(K)$ if and only if there exists a pair (z, w) such that (x, y) and (z, w)

together satisfy the extremality conditions. Such a pair (z, w) is a Kuhn-Tucker vector for $S(K)$.

PROOF. The equivalence is immediate from Lemma 5.10 and condition (e) of Theorem 5.11. From condition (a) of Theorem 5.11 it is not hard to show that (z, w) is a Kuhn-Tucker vector for $S(K)$.

By analogy with (36.6) for convex programming, one might ask whether the last assertion of Theorem 5.13 can be strengthened as follows: "For any given stable optimal solution (\bar{x}, \bar{y}) of $S(K)$, the pairs (z, w) such that (\bar{x}, \bar{y}) together with (z, w) satisfy the extremality conditions are precisely the Kuhn-Tucker vectors for $S(K)$." Unfortunately this fails in general, as Example 5.14 will show. There are conditions, however, under which this does hold. One such condition will now be given.

COROLLARY 5.13.1. Assume $S(K)$ has a unique stable optimal solution (\bar{x}, \bar{y}) and that $S(L)$ is strongly consistent. Then the Kuhn-Tucker vectors for $S(K)$ are precisely those pairs (z, w) such that (\bar{x}, \bar{y}) and (z, w) together satisfy the extremality conditions.

PROOF. By the theorem, if (\bar{x}, \bar{y}) and (z, w) together satisfy the extremality conditions then (z, w) is a Kuhn-Tucker vector for $S(K)$. Now suppose (z, w) is a Kuhn-Tucker vector for $S(K)$. Since $S(L)$ is strongly consistent, Corollary 5.5.1 implies that there exists some pair (x, y) such that (x, y) and (z, w) together satisfy the extremality conditions. By Theorem 5.13 this pair (x, y) is a stable optimal solution of $S(K)$. Hence the uniqueness assumption implies $(x, y) = (\bar{x}, \bar{y})$, and the proof is complete.

The reason that the strengthening of Theorem 5.13 proposed above fails to hold in general is that the set $\partial M^*(0, 0, 0, 0)$ of saddle points of the Lagrangian is the product set

$$\partial M^*(\cdot, \cdot, 0, 0)(0, 0) \times \partial M^*(0, 0, \cdot, \cdot)(0, 0) ,$$

which involves in each "factor" both the pair (x, y) of "solution variables" and the pair (z, w) of "Kuhn-Tucker variables." The following example illustrates this

EXAMPLE 5.14. Take $p = n$ and $q = m$ and define $S(K)$ by $K(u, x, v, y) = \langle u, y \rangle + \langle x, v \rangle$. It is easily checked that $A_i^{-1}\{0\} \cap (\text{rec cone}_i K) = \{(0, 0)\}$ for $i = 1, 2$. Hence Lemma 1.9 implies that $\text{range } A^* \cap \text{ri}(\text{dom } K^*) \neq \emptyset$, or in other words $S(L)$ is strongly consistent (see the proof of Lemma 5.4). If $[P]$ denotes the equivalence class containing P_1 and P_2 , then $[P] = [AK]$ by Theorem 5.5. Since

$$P_1(u, v) \begin{cases} 0 & \text{if } u = 0 \text{ and } v = 0 \\ +\infty & \text{if } u = 0 \text{ and } v \neq 0 \\ -\infty & \text{if } u \neq 0 , \end{cases}$$

this implies that $\text{dom } AK = \text{dom } P = \{(0, 0)\}$, whereas $A \text{ dom } K = \mathbb{R}^p \times \mathbb{R}^q$. Clearly $S(K)$ is strongly consistent, the set of optimal solutions of $S(K)$ is $\mathbb{R}^m \times \mathbb{R}^n$, and the set of Kuhn-Tucker vectors for $S(K)$ is $\mathbb{R}^p \times \mathbb{R}^q$. The Lagrangian of $S(K)$ contains the function

$$\begin{aligned} M(x, w, y, z) &= \sup_u \inf_v \{ \langle u, z \rangle + \langle v, w \rangle + K(u, x, v, y) \} \\ &= \begin{cases} 0 & \text{if } x + w = 0 \text{ and } y + z = 0 \\ +\infty & \text{if } x + w = 0 \text{ and } y + z \neq 0 \\ -\infty & \text{if } x + w \neq 0 . \end{cases} \end{aligned}$$

Hence the set of saddle points of the Lagrangian is just

$$\text{dom } M = \{(x, w) \mid x + w = 0\} \times \{(y, z) \mid y + z = 0\} .$$

Thus, if (z, w) is any given Kuhn-Tucker vector for $S(K)$, the set

$$\{(x, y) \mid (x, w, y, z) \text{ is a saddle point of the Lagrangian}\}$$

equals $\{(-w, -z)\}$ and hence is far from being equal to the set of optimal solutions of $S(K)$. It is perhaps of interest to note that in this example the dual program $S(L)$ is given by $L(s, z, t, w) = \langle s, w \rangle + \langle z, t \rangle$ and hence is "identical" with $S(K)$.

The next theorem says that stability of optimal solutions can often be assumed, inasmuch as the saddle programs considered will often be strongly consistent.

THEOREM 5.15. If $S(K)$ is strongly consistent, then every optimal solution of $S(K)$ is stable.

PROOF. Suppose $S(K)$ is strongly consistent, and let (x, y) be any optimal solution of $S(K)$. By the dual version of Corollary 5.5.1, there exists a pair (z, w) such that $(-x, 0, -y, 0) \in \partial L(0, z, 0, w)$. Hence Theorem 5.11 and Lemma 5.10 imply that (x, y) is stable.

This result can be combined with previous ones. Observe, for instance, that combining it with Theorem 5.13 yields an extension of the celebrated Kuhn-Tucker Theorem to generalized saddle programs (cf. (36.6)).

The following example shows that in the absence of strong consistency there may exist unstable optimal solutions.

EXAMPLE 5.16. In Example 5.3 take $m = n = 1$ and take $[J]$ to be such that $[J]^*$ contains the closed proper concave-convex function

$$(s, t) \rightarrow \begin{cases} -\sqrt{t} & \text{if } s \in [0, 1] \text{ and } t \in [0, 1] \\ +\infty & \text{if } s \in [0, 1] \text{ and } t \notin [0, 1] \\ -\infty & \text{if } s \notin [0, 1] . \end{cases}$$

Clearly $\text{dom } J^* = [0, 1] \times [0, 1]$, $\underline{J}^*(0, 0) = \overline{J}^*(0, 0)$, and $\partial J^*(0, 0) = \emptyset$. From the analysis in Example 5.3 it follows that $S(L)$ is consistent but fails to be strongly consistent, $S(L)$ has a well-defined primal problem, and $(0, 0)$ is the only optimal solution of $S(L)$. If $(0, 0)$ were stable, then by Theorem 5.12 there would exist a stable optimal solution of $S(K)$. But the set of optimal solutions of $S(K)$ is easily seen to be $\partial J^*(0, 0)$, which is empty. Hence $(0, 0)$ is an unstable optimal solution of $S(L)$.

This example also shows something else. Notice that by Corollary 5.5.1 and Theorem 5.11, if the dual program $S(L)$ is strongly consistent then each Kuhn-Tucker vector (z, w) for $S(K)$ "corresponds" to a saddle point of the Lagrangian in the sense that $(0, 0, 0, 0) \in \partial M(x, w, y, z)$ for some pair (x, y) . Example 5.16 shows that this need not be true when $S(L)$ fails to be strongly consistent.

§6: Ordinary Saddle Programs and Lagrange Multipliers

In this section we shall turn our attention to minimax problems constrained by finitely many concave and convex inequalities. Such problems will be cast in the framework of generalized saddle programs. By applying the general theory of §5, we shall obtain information regarding the Lagrange multiplier principle for such problems and also an explicit description of a dual minimax problem.

Suppose that S and T are nonempty closed convex subsets of R^m and R^n , respectively, and that H is a finite continuous concave-convex function on $S \times T$, g_1, \dots, g_p are finite upper semi-continuous concave functions on S , and f_1, \dots, f_q are finite lower semi-continuous convex functions on T . The problem we consider here is that of finding the saddle points of H with respect to the pairs (x, y) in $S \times T$ satisfying the constraints

$$g_1(x) \geq 0, \dots, g_p(x) \geq 0$$

and

$$f_1(y) \leq 0, \dots, f_q(y) \leq 0.$$

Actually, though, by treating this problem within the framework of generalized saddle programs we are able to deal with the following more general situation. Let H be a closed proper concave-convex function on $R^m \times R^n$, let each g_1, \dots, g_p be a closed proper concave function on R^m satisfying

$$\text{dom}_1 H \subset \text{dom } g_1 \text{ and } \text{ri}(\text{dom}_1 H) \subset \text{ri}(\text{dom } g_1),$$

and let each f_1, \dots, f_q be a closed proper convex function on R^n satisfying

$$\text{dom}_2 H \subset \text{dom } f_j \text{ and } \text{ri}(\text{dom}_2 H) \subset \text{ri}(\text{dom } f_j) .$$

The problem now is to find the saddle points of H with respect to the pairs (x, y) in $R^m \times R^n$ satisfying the above inequality constraints. This situation is indeed more general. Just take H to be the lower simple extension to all of $R^m \times R^n$ of the previous function on $S \times T$, extend the g_i 's to be $-\infty$ on $R^m \setminus S$, and extend the f_j 's to be $+\infty$ on $R^n \setminus T$. Then $\text{dom } H = S \times T$, and the effective domain inclusions above are automatically satisfied.

To obtain a generalized saddle program we define subsets $C \subset R^p \times R^m$ and $D \subset R^q \times R^n$ by

$$C = \{(u, x) \mid x \in \text{dom}_1 H, \quad g_1(x) \geq u_1, \dots, g_p(x) \geq u_p\}$$

and

$$D = \{(v, y) \mid y \in \text{dom}_2 H, \quad f_1(y) \leq v_1, \dots, f_q(y) \leq v_q\} ,$$

and let K be the function on $(R^p \times R^m) \times (R^q \times R^n)$ defined by

$$K(u, x, v, y) = \begin{cases} H(x, y) & \text{if } (u, x) \in C \text{ and } (v, y) \in D \\ +\infty & \text{if } (u, x) \in C \text{ and } (v, y) \notin D \\ -\infty & \text{if } (u, x) \notin C . \end{cases}$$

Our first theorem shows that this K does determine a generalized saddle program.

THEOREM 6.1. The function K is closed proper concave-convex with effective domain $C \times D$. Moreover,

$$\text{ri } C = \{(u, x) \mid x \in \text{ri}(\text{dom}_1 H) \text{ and } g_1(x) > u_1, \dots, g_p(x) > u_p\}$$

and

$$\text{cl } C = \{(u, x) \mid x \in \text{cl}(\text{dom}_1 H) \text{ and } g_1(x) \geq u_1, \dots, g_p(x) \geq u_p\},$$

and similar formulas hold for $\text{ri } D$ and $\text{cl } D$.

PROOF. Define functions H_0, \dots, H_{p+q} on $(R^p \times R^m) \times (R^q \times R^n)$ as follows:

$$H_0(u, x, v, y) = H(x, y)$$

$$H_1(u, x, v, y) = \begin{cases} 0 & \text{if } (x, u_1) \in \text{epi } g_1 \\ -\infty & \text{if } (x, u_1) \notin \text{epi } g_1 \end{cases} \quad l = 1, \dots, p$$

$$H_{p+j}(u, x, v, y) = \begin{cases} 0 & \text{if } (y, v_j) \in \text{epi } f_j \\ +\infty & \text{if } (y, v_j) \notin \text{epi } f_j \end{cases} \quad j = 1, \dots, q.$$

Clearly,

$$\text{dom } H_0 = (R^p \times \text{dom}_1 H) \times (R^q \times \text{dom}_2 H)$$

$$\text{dom } H_l = \{(u, x) \mid (x, u_l) \in \text{epi } g_l\} \times (R^q \times R^n) \quad l = 1, \dots, p$$

$$\text{dom } H_{p+j} = (R^p \times R^m) \times \{(v, y) \mid (y, v_j) \in \text{epi } f_j\} \quad j = 1, \dots, q$$

and from (34.3) it follows that each H_k is closed and proper. Since

$$\text{ri}(\text{dom } H_0) \cap \dots \cap \text{ri}(\text{dom } H_{p+q}) \neq \emptyset,$$

Theorem 3.2 implies that $[H_0] + \dots + [H_{p+q}]$ is well-defined, has effective domain

$$C \times D = \text{dom } H_0 \cap \dots \cap \text{dom } H_{p+q},$$

and contains the function K . The formulas for $\text{ri } C$ and $\text{cl } C$ (resp.

$\text{ri } D$ and $\text{cl } D$) follow from (6.5), (7.3) and the fact that $\text{epi } g_l$ (resp.

$\text{epi } f_j$) is closed.

According to the theorem, $[K]$ is a generalized saddle program $S(K)$ on $R^m \times R^n$ with perturbations in $R^p \times R^q$. We call it the ordinary saddle program associated with $H, g_1, \dots, g_p, f_1, \dots, f_q$.

It will be convenient to introduce the following notation. For any subset S of $R^p \times R^m$ write $S_u = \{x \mid (u, x) \in S\}$ for each $u \in R^p$. Similarly, for any subset T of $R^q \times R^n$ write $T_v = \{y \mid (v, y) \in T\}$ for each $v \in R^q$.

Since the feasible solutions of any generalized saddle program are those pairs (x, y) such that $(0, x, 0, y) \in \text{dom } K$, the set of feasible solutions of the ordinary saddle program $S(K)$ is just $C_0 \times D_0$, i.e.

$$\{(x, y) \in \text{dom } H \mid g_1(x) \geq 0, \dots, g_p(x) \geq 0 \text{ and } f_1(y) \leq 0, \dots, f_q(y) \leq 0\}.$$

Recall from the general theory that $S(K)$ is consistent if and only if $S(K)$ has a feasible solution, i.e. if and only if $C_0 \times D_0$ is nonempty.

According to the following corollary, strong consistency of $S(K)$ is analogous to the Slater constraint qualification encountered in convex programming.

COROLLARY 6.1.1. The program $S(K)$ is strongly consistent if and only if there exists a pair (x, y) in $\text{dom } H$ such that $g_1(x) > 0, \dots, g_p(x) > 0$ and $f_1(y) < 0, \dots, f_q(y) < 0$.

PROOF. By the formulas for $\text{ri } C$ and $\text{ri } D$ given in Theorem 6.1, the assertion of the corollary holds with $\text{dom } H$ replaced by $\text{ri}(\text{dom } H)$.

Now suppose $(x, y) \in \text{dom } H$ is such that $g_1(x) > 0, \dots, g_p(x) > 0$ and $f_1(y) < 0, \dots, f_q(y) < 0$. Let (x_1, y_1) be any element of $\text{ri}(\text{dom } H)$. Then

(6.1) and (7.5) imply that, for sufficiently small positive λ , the pair

$$(x_\lambda, y_\lambda) = (1 - \lambda)(x, y) + \lambda(x_1, y_1)$$

is in $\text{ri}(\text{dom } H)$ and satisfies $g_1(x_\lambda) > 0, \dots, g_p(x_\lambda) > 0$ and $f_1(y_\lambda) < 0, \dots, f_q(y_\lambda) < 0$.

If $S(K)$ is strongly consistent, then all the objective functions of $S(K)$ are equivalent and hence the notions of optimal value and optimal solution can be expressed in terms of the single objective function

$$K(0, x, 0, y) = \begin{cases} H(x, y) & \text{if } x \in C_0 \text{ and } y \in D_0 \\ +\infty & \text{if } x \in C_0 \text{ and } y \notin D_0 \\ -\infty & \text{if } x \notin C_0 \end{cases}$$

In this event, the optimal value in $S(K)$ exists and equals α if and only if

$$\sup_{C_0} \inf_{D_0} H = \inf_{D_0} \sup_{C_0} H = \alpha \in \mathbb{R},$$

and (x, y) is an optimal solution of $S(K)$ if and only if (x, y) is a saddle point of H with respect to $C_0 \times D_0$. Thus the ordinary saddle program $S(K)$ accurately reflects the original minimax problem we set out to study. Even when $S(K)$ is not strongly consistent, the relationship it bears to the original minimax problem is only slightly more complicated, as explained in the next corollary.

COROLLARY 6.1.2. Write $\text{cl } C = \bar{C}$ and $\text{cl } D = \bar{D}$. The optimal value
in $S(K)$ exists and equals α if and only if

$$\sup_{C_0} \inf_{\bar{D}_0} \underline{H} = \inf_{D_0} \sup_{\bar{C}_0} \bar{H} = \alpha \in \mathbb{R} .$$

A pair (x, y) is an optimal solution of $S(K)$ if and only if

$$\inf_{\bar{D}_0} \underline{H}(x, \cdot) = \sup_{\bar{C}_0} \bar{H}(\cdot, y) \in \mathbb{R} .$$

A pair (z, w) is a Kuhn-Tucker vector for $S(K)$ if and only if the optimal
value in $S(K)$ exists and equals α and

$$\langle u, z \rangle + \inf_{D_0} \sup_{\bar{C}_u} \bar{H} \leq \alpha \leq \sup_{C_0} \inf_{\bar{D}_v} \underline{H} + \langle v, w \rangle, \quad \forall u \quad \forall v .$$

An optimal solution (x, y) of $S(K)$ is stable if and only if

$$\lim_{\lambda \downarrow 0} \frac{1}{\lambda} \left\{ \inf_{\bar{D}_{\lambda v}} \underline{H}(x, \cdot) - \inf_{\bar{D}_0} \underline{H}(x, \cdot) \right\} > -\infty, \quad \forall v$$

and

$$\lim_{\lambda \downarrow 0} \frac{1}{\lambda} \left\{ \sup_{\bar{C}_{\lambda u}} \bar{H}(\cdot, y) - \sup_{\bar{C}_0} \bar{H}(\cdot, y) \right\} < +\infty, \quad \forall u .$$

PROOF. By Theorems 6.1 and 0.1, the least member \underline{K} of $[K]$ is
the convex closure of K . Direct computation thus yields

$$\underline{K}(u, x, v, y) = \begin{cases} \underline{H}(x, y) & \text{if } (u, x) \in C \text{ and } (v, y) \in \bar{D} \\ +\infty & \text{if } (u, x) \in C \text{ and } (v, y) \notin \bar{D} \\ -\infty & \text{if } (u, x) \notin C \end{cases}$$

and hence

$$P_1(u, v) = \sup_{C_u} \inf_{\bar{D}_v} H.$$

Analogous formulas hold for \bar{K} and P_2 . A pair (z, w) is a Kuhn-Tucker vector for $S(K)$ if and only if $P_1(0, 0) = P_2(0, 0) = \alpha \in R$ and $\langle u, z \rangle + P_2(u, 0) \leq \alpha \leq P_1(0, v) + \langle v, w \rangle$, $\forall u \forall v$. Since $P_1(0, 0) = P_2(0, 0) = \alpha \in R$ occurs if and only if the optimal value in $S(K)$ exists and equals α , the first two assertions of the corollary follow immediately from the formulas for P_1 and P_2 . Now recall from §5 the functions $f_x(v) = \inf_{\bar{D}_v} \underline{K}(0, x, v, \cdot)$ and $g_y(u) = \sup_{\bar{C}_u} \bar{K}(u, \cdot, 0, y)$. The formulas for \underline{K} and \bar{K} imply that

$$f_x(v) = \inf_{\bar{D}_v} \underline{H}(x, \cdot)$$

whenever $x \in C_0$ and

$$g_y(u) = \sup_{\bar{C}_u} \bar{H}(\cdot, y)$$

whenever $y \in D_0$. Since a pair (x, y) is an optimal solution of $S(K)$ if and only if $f_x(0) = g_y(0) \in R$ (and $(x, y) \in C_0 \times D_0$), the third assertion follows immediately. The last assertion also follows immediately from the formulas for f_x and g_y .

All the general theory of §5 can be applied to the ordinary saddle program $S(K)$. Particularly strong results can be obtained under the hypothesis that $S(K)$ is strongly consistent and $\text{dom } K_0$ is bounded (i.e., that the set of feasible solutions is bounded and that some feasible solution satisfies the inequality constraints with strict inequality). In what follows, though, we shall be concerned principally with the question of whether there exists a "good" Lagrange multiplier principle for ordinary saddle programs. Then at the end of the section we shall identify the dual of an ordinary saddle program and explicitly describe the dual minimax problem.

The next result shows that the Lagrangian of $S(K)$, which we obtain via the general theory of §5, looks like what one would expect. Also, the corollary which follows shows that the extremality conditions we obtain via §5 are direct analogues of the familiar Kuhn-Tucker conditions of convex programming.

THEOREM 6.2. The Lagrangian of $S(K)$ contains the function

$$(x, w, y, z) \rightarrow \begin{cases} H(x, y) + \sum_i z_i g_i(x) + \sum_j w_j f_j(y) & \text{if } (x, w) \in S \text{ and } (y, z) \in T \\ +\infty & \text{if } (x, w) \in S \text{ and } (y, z) \notin T \\ -\infty & \text{if } (x, w) \notin S, \end{cases}$$

where $S = \text{dom}_1 H \times R_+^q$ and $T = \text{dom}_2 H \times R_+^p$.

PROOF. By definition, the Lagrangian contains the function

$$\begin{aligned}
M(x, w, y, z) &= \sup_u \inf_v \{ \langle u, z \rangle + \langle v, w \rangle + K(u, x, v, y) \} \\
&= \sup_{C_x} \inf_{D_y} \{ \langle u, z \rangle + \langle v, w \rangle + H(x, y) \} ,
\end{aligned}$$

where $C_x = \{u \mid (u, x) \in C\}$ and $D_y = \{v \mid (v, y) \in D\}$. Now C_x equals $\{u \mid g_1(x) \geq u_1, \dots, g_p(x) \geq u_p\}$ when $x \in \text{dom}_1 H$ and equals the empty set otherwise, and similarly D_y equals $\{v \mid f_1(y) \leq v_1, \dots, f_q(y) \leq v_q\}$ when $y \in \text{dom}_2 H$ and equals the empty set otherwise. Therefore the conventions imply that $M(x, w, y, z) = -\infty$ when $x \notin \text{dom}_1 H$ and $M(x, w, y, z) = +\infty$ when $x \in \text{dom}_1 H$ and $y \notin \text{dom}_2 H$. When $(x, y) \in \text{dom } H$,

$$\begin{aligned}
M(x, w, y, z) &= H(x, y) + \sup_{C_x} \{ \langle u, z \rangle + \inf_{D_y} \{ \langle v, w \rangle \} \} \\
&= \begin{cases} -\infty & \text{if } w \notin R_+^q \\ +\infty & \text{if } w \in R_+^q \text{ and } z \notin R_+^p \\ H(x, y) + \sum z_i g_i(x) + \sum w_j f_j(y) & \text{if } w \in R_+^q \text{ and } z \in R_+^p . \end{cases}
\end{aligned}$$

It is easy to show that $\underline{M} = \text{cl}_2 M$ is given by

$$\underline{M}(x, w, y, z) = \begin{cases} \underline{H}(x, y) + \sum z_i g_i(x) + \sum w_j f_j(y) & \text{if } (x, w) \in S \text{ and } (y, z) \in \text{cl } T \\ +\infty & \text{if } (x, w) \in S \text{ and } (y, z) \notin \text{cl } T \\ -\infty & \text{if } (x, w) \notin S . \end{cases}$$

Finally, observe that the function in the theorem is bounded below by \underline{M} and above by M .

COROLLARY 6.2.1. Two pairs $(x, y) \in R^m \times R^n$ and $(z, w) \in R^p \times R^q$ satisfy the extremality conditions associated with $S(K)$ if and only if

$$(x, y) \in C_0 \times D_0, \quad (z, w) \in R_+^p \times R_+^q,$$

$$z_i g_i(x) = 0 \quad \text{for } i = 1, \dots, p,$$

$$w_j f_j(y) = 0 \quad \text{for } j = 1, \dots, q,$$

$$0 \in \partial_1 H(x, y) + \sum \partial(z_i g_i)(x),$$

$$0 \in \partial_2 H(x, y) + \sum \partial(w_j f_j)(y).$$

The term $\sum \partial(z_i g_i)(x)$ can be replaced by $\sum z_i \partial g_i(x)$ if the summation is taken only over those i such that $z_i > 0$. Similarly, the term $\sum \partial(w_j f_j)(y)$ can be replaced by $\sum w_j \partial f_j(y)$ if the summation is taken only over those j such that $w_j > 0$.

PROOF. By definition, (x, y) and (z, w) satisfy the extremality conditions if and only if (x, w, y, z) is a saddle point of the Lagrangian. By the theorem and (36.3), this occurs if and only if $(x, y) \in \text{dom } H$ and $(z, w) \in R_+^p \times R_+^q$,

$$H(x', y) + \sum z_i g_i(x') + \sum w_j f_j(y) \leq H(x, y) + \sum z_i g_i(x) + \sum w_j f_j(y) \quad (1)$$

for all $(x', w') \in \text{dom}_1 H \times R_+^q$, and

$$H(x, y) + \sum z_i g_i(x) + \sum w_j f_j(y) \leq H(x, y') + \sum z_i g_i(x) + \sum w_j f_j(y') \quad (2)$$

for all $(y', z') \in \text{dom}_2 H \times R_+^p$. Taking $z' = z$ in (2) and using (23.8)

implies $0 \in \partial_2 H(x, y) + \sum \partial(w_j f_j)(y)$. Taking $y' = y$, $z'_1 = 1 + z_1$ and

$z'_k = z_k$ for $k \neq 1$, (2) implies that $0 \leq g_1(x)$. This holds for each i .

But taking $y' = y$ and $z' = 0$ in (2) implies $\sum z_i g_i(x) \leq 0$. Hence

$z_i g_i(x) = 0$ for each i . Similarly, (1) implies that $f_j(y) \leq 0$ and $w_j f_j(y) = 0$

for each j and $0 \in \partial_1 H(x, y) + \sum \partial(z_i g_i)(x)$. This establishes one implication,

and the converse is now clear. Now observe that $w_j > 0$ trivially

implies $\partial(w_j f_j)(y) = w_j \partial f_j(y)$. On the other hand, if $w_j = 0$ then $y \in \text{dom}_2 H \subset \text{dom } f_j$ implies $\partial(w_j f_j)(y) = \partial \delta(y | \text{dom } f_j) \subset \partial \delta(y | \text{dom}_2 H) = 0^+ \partial_2 H(x, y)$ and hence $\partial_2 H(x, y) + \partial(w_j f_j)(y) = \partial_2 H(x, y)$. Thus the term $\Sigma \partial(w_j f_j)(y)$ can be replaced as indicated. The other assertion is proved similarly.

Variables of the sort z_1, \dots, z_p and w_1, \dots, w_q appearing in the Lagrangian of $S(K)$ are known traditionally as Lagrange multipliers. Sometimes this term also denotes the particular values of these variables which satisfy certain "extremality conditions" relating to a "Lagrangian function." In this second sense, Lagrange multipliers $(z_1, \dots, z_p, w_1, \dots, w_q) = (z, w)$ for an ordinary or generalized saddle program necessarily form a Kuhn-Tucker vector for the program (Theorem 5.13). However, a Kuhn-Tucker vector need not satisfy the extremality conditions, i.e. need not be a Lagrange multiplier. (This behavior can occur if the dual program fails to be strongly consistent. See the remarks following Example 5.16.) In other words, Kuhn-Tucker vectors are defined even when the extremality conditions are not satisfiable. Thus, Kuhn-Tucker vectors (rather than Lagrange multipliers) are the natural "equilibrium price vectors" for regularized saddle point problems.

By the general theory of §5, if (x, y) and (z, w) satisfy the extremality conditions then the optimal value in $S(K)$ exists and equals $H(x, y)$, (x, y) is a stable optimal solution of $S(K)$, and (z, w) is a Kuhn-Tucker vector for $S(K)$. In fact, such pairs (x, y) and (z, w) actually satisfy

$$\langle u, z \rangle + \bar{H}(x', y) \leq H(x, y) \leq \underline{H}(x, y') + \langle v, w \rangle$$

for every $(u, x') \in \text{cl } C$ and every $(v, y') \in \text{cl } D$. (cf. Corollary 6.1.2).

The following result can be viewed as the main existence theorem for "Lagrange multipliers." Notice that the conditions it gives are satisfied, for example, whenever

$$0^+ \text{dom}_1 H \cap \text{rec cone } g_1 \cap \dots \cap \text{rec cone } g_p = \{0\}$$

and

$$0^+ \text{dom}_2 H \cap \text{rec cone } f_1 \cap \dots \cap \text{rec cone } f_q = \{0\}.$$

THEOREM 6.3. If $S(K)$ is strongly consistent, then the extremality conditions can be satisfied whenever the sets

$$\{x \in \bigcap_{i=1}^p \text{rec cone } g_i \mid \inf \{\text{rec } H(\cdot, y)(x) \mid (0, y) \in \text{ri } D\} \geq 0\}$$

and

$$\{y \in \bigcap_{j=1}^q \text{rec cone } f_j \mid \sup \{\text{rec } H(x, \cdot)(y) \mid (0, x) \in \text{ri } C\} \leq 0\}$$

are closed under scalar multiplication by -1 .

PROOF. By Theorem 5.15 and Theorem 5.13, if $S(K)$ is strongly consistent and has an optimal solution then the extremality conditions can be satisfied. The remainder of the proof consists of showing that Corollary 5.2.1 applies to yield an optimal solution. Suppose $S(K)$ is strongly consistent. By Theorem 5.2, $S(K)$ has a well-defined primal problem which is given by the closed proper equivalence class $[K_0]$. Moreover, $\text{dom } K_0 = C_0 \times D_0$, $\text{ri}(\text{dom } K_0) = (\text{ri } C)_0 \times (\text{ri } D)_0$, and $[K_0]$ contains the function

$$K_0(x, y) = \begin{cases} H(x, y) & \text{if } x \in C_0 \text{ and } y \in D_0 \\ +\infty & \text{if } x \in C_0 \text{ and } y \notin D_0 \\ -\infty & \text{if } x \notin C . \end{cases}$$

Let $Y = \{y \mid f_1(y) \leq 0, \dots, f_q(y) \leq 0\}$. Then $D_0 = Y \cap \text{dom}_2 H$, so that $K_0(x, \cdot) = H(x, \cdot) + \delta(\cdot \mid Y)$ whenever $x \in \text{ri}(\text{dom}_1 H)$. It follows from the definitions and (9.3) that

$$(\text{rec}_2 K_0)(y) = \sup \{ \text{rec } H(x, \cdot)(y) + \text{rec } \delta(\cdot \mid Y)(y) \mid (0, x) \in \text{ri } C \} .$$

Now note that $\text{rec } \delta(\cdot \mid Y) = \delta(\cdot \mid 0^+ Y)$ by (8.5), and $0^+ Y = \bigcap_{j=1}^q \text{rec cone } f_j$ by (8.3.3) and (8.7). These facts together imply that $(\text{rec}_2 K_0)(y) \leq 0$ if and only if $y \in \bigcap_{j=1}^q \text{rec cone } f_j$ and $\sup \{ \text{rec } H(x, \cdot)(y) \mid (0, x) \in \text{ri } C \} \leq 0$.

A similar argument shows that $(\text{rec}_1 K_0)(x) \geq 0$ if and only if $x \in \bigcap_{i=1}^p \text{rec cone } g_i$ and $\inf \{ \text{rec } H(\cdot, y)(x) \mid (0, y) \in \text{ri } D \} \geq 0$. These two equivalences show that the hypothesis is just what is needed to apply Corollary 5.2.1.

For our results concerning the Lagrange multiplier principle we need certain functions $H_{z, w}$ and sets $S_{z, w}$. For each $(z, w) \in R_+^p \times R_+^q$ define a function $H_{z, w}$ on $R^m \times R^n$ by

$$H_{z, w}(x, y) = \begin{cases} H(x, y) + \sum z_i g_i(x) + \sum w_j f_j(y) & \text{if } x \in \text{dom}_1 H \text{ and } y \in \text{dom}_2 H \\ +\infty & \text{if } x \in \text{dom}_1 H \text{ and } y \notin \text{dom}_2 H \\ -\infty & \text{if } x \notin \text{dom}_1 H . \end{cases}$$

By Theorem 3.2 it follows easily from the blanket regularity assumptions that $H_{z, w}$ is closed and proper and has the same effective domain as H .

Observe that if \tilde{M} denotes the particular Lagrangian given in Theorem 6.2,

$$H_{z,w}(x,y) = \tilde{M}(x,w,y,z)$$

for every $(z,w) \in R_+^p \times R_+^q$ and every $(x,y) \in R^m \times R^n$. If $(z,w) \notin R_+^p \times R_+^q$ put $S_{z,w} = \emptyset$, and if $(z,w) \in R_+^p \times R_+^q$ let $S_{z,w}$ denote the set of pairs (x,y) which are saddle points of $H_{z,w}$ and which satisfy the conditions

$$g_i(x) \geq 0 \text{ and } z_i g_i(x) = 0 \quad \text{for } i = 1, \dots, p$$

and

$$f_j(y) \leq 0 \text{ and } w_j f_j(y) = 0 \quad \text{for } j = 1, \dots, q.$$

(These latter conditions together with the condition $(z,w) \in R_+^p \times R_+^q$ are traditionally called complementary slackness conditions.)

For ordinary convex programs there exists a good Lagrange multiplier principle (Theorem 28.1 in [48]). The analogous result for ordinary saddle programs would be the following: "If (z,w) is any Kuhn-Tucker vector for $S(K)$, then $S_{z,w}$ is precisely the set of optimal solutions of $S(K)$." As we shall see in Theorem 6.5, however, the situation is in general more complicated than this. To describe the situation fully, we need to know the connection between the sets $S_{z,w}$ and the extremality conditions.

LEMMA 6.4. The pairs $(x,y) \in R^m \times R^n$ and $(z,w) \in R_+^p \times R_+^q$

together satisfy the extremality conditions if and only if $(x,y) \in S_{z,w}$.

PROOF. For any $(z,w) \in R_+^p \times R_+^q$, a pair (x,y) is a saddle point of $H_{z,w}$ if and only if $(x,y) \in \text{dom } H$,

$$H(x',y) + \sum z_i g_i(x') \leq H(x,y) + \sum z_i g_i(x), \quad \forall x' \in \text{dom}_1 H$$

and

$$H(x, y) + \sum w_j f_j(y) \leq H(x, y') + \sum w_j f_j(y'), \quad \forall y' \in \text{dom}_2 H.$$

Now it is an easy exercise to show (using (7.5) and (6.1)) that for any convex function f and any convex set C containing $\text{ri}(\text{dom } f)$, $x^* \in \partial f(x)$ if and only if

$$f(x') \geq f(x) + \langle x^*, x' - x \rangle, \quad \forall x' \in C.$$

But $\text{ri}(\text{dom}(H(x, \cdot) + \sum w_j f_j)) = \text{ri}(\text{dom}_2 H)$ and $\text{ri}(\text{dom}(H(\cdot, y) + \sum z_i g_i)) = \text{ri}(\text{dom}_1 H)$ whenever $(z, w) \in R_+^p \times R_+^q$ and $(x, y) \in \text{dom } H$. Hence it follows from these facts and (23.8) that for $(z, w) \in R_+^p \times R_+^q$, (x, y) is a saddle point of $H_{z, w}$ if and only if $(x, y) \in \text{dom } H$,

$$0 \in \partial_1 H(x, y) + \sum \partial(z_i g_i)(x)$$

and

$$0 \in \partial_2 H(x, y) + \sum \partial(w_j f_j)(y).$$

The lemma follows trivially from this by Corollary 6.2.1 and the definition of $S_{z, w}$.

We can now present the theorem promised. Notice that the first part of it constitutes an extension of the Kuhn-Tucker Theorem to ordinary saddle programs.

THEOREM 6.5. The set of stable optimal solutions of $S(K)$ is precisely

$$\{S_{z,w} \mid (z,w) \in R^p \times R^q\},$$

and when $S(K)$ is strongly consistent this set coincides with the set of all optimal solutions of $S(K)$. If $S_{z,w} \neq \emptyset$, then (z,w) is a Kuhn-Tucker vector for $S(K)$; the converse holds when the program dual to $S(K)$ is strongly consistent.

PROOF. In view of Lemma 6.4, the first assertion is immediate from Theorems 5.13 and 5.15, and the second assertion is immediate from Theorems 5.13 and 5.11 together with Corollary 5.5.1

From Theorem 6.5 we can conclude that in general there does not exist a good Lagrange multiplier principle for ordinary saddle programs, i. e. "good" in the sense that a single Kuhn-Tucker vector (z,w) generates all the (stable) optimal solutions via $S_{z,w}$. But, one could ask, might it not be possible to obtain a good Lagrange multiplier principle by recasting the given constrained minimax problem as another generalized saddle program involving a class of perturbations different from that which we actually used? Concerning this, observe that any "Lagrange multiplier principle" by its very nature would involve a Lagrangian function of the sort given in Theorem 6.2. And as pointed out in §5, such a Lagrangian uniquely determines the dual pair of generalized saddle programs and hence the particular classes of perturbations. This means that the perturbations introduced at the beginning of §6 are in fact the only ones relevant to the issue of a Lagrange multiplier principle. Hence our conclusion on the

basis of Theorem 6.5 that no good Lagrange multiplier principle exists in general for ordinary saddle programs.

In certain circumstances, however, there is an analogue of (28.1) for ordinary saddle programs.

COROLLARY 6.5.1. Assume that the program dual to $S(K)$ is strongly consistent. If the pair (z, w) is a unique Kuhn-Tucker vector for $S(K)$, or equivalently if (z, w) is a unique optimal solution of the dual program, then the set of stable optimal solutions of $S(K)$ is nonempty and equals $S_{z, w}$.

PROOF. Immediate from the theorem and Corollary 5.5.1.

Theorem 6.5 and Corollary 6.5.1 are actually valid for any generalized saddle program, provided that for each $(z, w) \in R^p \times R^q$ the set $S_{z, w}$ is redefined to be the set of pairs $(x, y) \in R^m \times R^n$ such that (x, y) and (z, w) together satisfy the extremality conditions. The proofs go through exactly the same except that this new definition of $S_{z, w}$ plays the role of Lemma 6.4.

We should emphasize that the intent of the discussion following Theorem 6.5 was not to say that the method of Lagrange multipliers cannot be used to advantage in ordinary saddle programming. On the contrary, the only point being made was that one cannot in general expect to obtain all of the (stable) optimal solutions of $S(K)$ via a single $S_{z, w}$.

To implement the method of Lagrange multipliers, according to Theorem 6.5 we need to find a pair (z, w) such that $S_{z, w}$ is nonempty. Now Theorem 6.5 together with Corollary 5.5.1 show that, in the presence of dual strong consistency, the pairs (z, w) for which $S_{z, w}$ is nonempty

are precisely the optimal solutions of the dual program $S(L)$, i.e., the saddle points of the dual minimax problem given by $[L_0]$. In view of this, we shall devote the rest of this section to a basic description of $S(L)$ and $[L_0]$.

We begin by giving a (primal) characterization of dual strong consistency. Notice that the conditions given are satisfied, for example, whenever

$$0^+ \text{ dom}_1 H \cap \text{rec cone } g_1 \cap \dots \cap \text{rec cone } g_p = \{0\}$$

and

$$0^+ \text{ dom}_2 H \cap \text{rec cone } f_1 \cap \dots \cap \text{rec cone } f_q = \{0\}.$$

LEMMA 6.6. The program dual to $S(K)$ is strongly consistent if and only if the two sets

$$\text{rec cone}_1 H \cap \text{rec cone } g_1 \cap \dots \cap \text{rec cone } g_p,$$

$$\text{rec cone}_2 H \cap \text{rec cone } f_1 \cap \dots \cap \text{rec cone } f_q$$

are closed under scalar multiplication by -1 .

PROOF. By Lemma 5.4 it suffices to show that

$$(\text{rec}_1 K)(0, x) \geq 0 \text{ if and only if } x \in \text{rec cone}_1 H \cap \bigcap_{i=1}^p \text{rec cone } g_i$$

and

$$(\text{rec}_2 K)(0, y) \leq 0 \text{ if and only if } y \in \text{rec cone}_2 H \cap \bigcap_{j=1}^q \text{rec cone } f_j.$$

Let H_0, \dots, H_{p+q} be as in the proof of Theorem 6.1 and let $(u, x) \in \text{ri } C$.

By (9.3),

$$\text{rec } K(u, x, \cdot, \cdot) = \sum \text{rec } H_k(u, x, \cdot, \cdot) .$$

Observe that trivially

$$\text{rec } H_0(u, x, \cdot, \cdot)(v, y) = \text{rec } H(x, \cdot)(y)$$

and

$$\text{rec } H_i(u, x, \cdot, \cdot)(v, y) = 0 \text{ for } i = 1, \dots, p .$$

With the aid of (8.5) and (9.1) it is easy to show that

$$\text{rec } H_{p+j}(u, x, \cdot, \cdot)(v, y) = \delta((y, v_j) | \text{epi}(\text{rec } f_j))$$

for $j = 1, \dots, q$. These facts together imply that $(\text{rec}_2 K)(v, y)$ equals $(\text{rec}_2 H)(y)$ when $(\text{rec } f_j)(y) \leq v_j$ for $j = 1, \dots, q$ and equals $+\infty$ otherwise. This establishes the second equivalence stated above. The first can be proved similarly.

The next result gives explicit representatives of the program $S(L)$ dual to the ordinary saddle program $S(K)$.

THEOREM 6.7. The function

$$(s, z, t, w) \rightarrow \begin{cases} \sup_{x \in C_H} \inf_{y \in D_H} \{ \langle x, -s \rangle + \langle y, -t \rangle + H_{z, w}(x, y) \} & \text{if } z \in R_+^p \text{ and } w \in R_+^q \\ -\infty & \text{if } z \in R_+^p \text{ and } w \notin R_+^q \\ +\infty & \text{if } z \notin R_+^p \end{cases} ,$$

where $C_H \times D_H = \text{dom } H$, belongs to $[L]$. The function obtained from this formula by interchanging $\sup_{x \in C_H}$ with $\inf_{y \in D_H}$ also belongs to $[L]$.

PROOF. Consider the function

$$(x, w, y, z) \rightarrow \begin{cases} H_{z, w}(x, y) & \text{if } w \in R_+^q \text{ and } z \in R_+^p \\ -\infty & \text{if } w \notin R_+^q \text{ and } z \in R_+^p \\ +\infty & \text{if } z \notin R_+^p \end{cases}$$

It can be checked easily that this function is concave-convex and also (using the formula for \underline{M} given in the proof of Theorem 6.2, plus the parallel formula for \bar{M}) that it is bounded below by \underline{M} and above by \bar{M} . Hence this function is a member of the Lagrangian equivalence class $[M]$. It follows from Theorem 5.8 and the definition of the partial conjugacy operation that $[L]$ contains both the function

$$(s, z, t, w) \rightarrow \begin{cases} \sup_x \inf_y \{ \langle x, -s \rangle + \langle y, -t \rangle + H_{z, w}(x, y) \} & \text{if } z \in R_+^p \text{ and } w \in R_+^q \\ -\infty & \text{if } z \in R_+^p \text{ and } w \notin R_+^q \\ +\infty & \text{if } z \notin R_+^p \end{cases}$$

and the analogous function obtained by interchanging \sup with \inf . Finally, the definition of $H_{z, w}$ allows us to restrict the x 's to C_H and the y 's to D_H .

Using the elements of $[L]$ just given, we can characterize $\text{dom } L$. This furnishes an alternate, direct way of checking for dual strong consistency (cf. the primal characterization given in Lemma 6.6).

THEOREM 6.8. Let $s \in R^m$, $t \in R^n$, $z \in R^p$ and $w \in R^q$. Then

$$(s, z) \in \text{dom}_1 L \iff z \in R_+^p \text{ and } s \in \text{dom}_1 H^* + \sum z_i \text{dom}_1 g_i^*$$

and

$$(t, w) \in \text{dom}_2 L \iff w \in R_+^q \text{ and } t \in \text{dom}_2 H^* + \sum w_j \text{dom} f_j^*.$$

PROOF. For any $\tilde{L} \in [L]$, $(t, w) \in \text{dom}_2 L$ if and only if $\tilde{L}(\cdot, \cdot, t, w)$ is never $-\infty$. Taking for \tilde{L} the function displayed in Theorem 6.7 yields $(t, w) \in \text{dom}_2 L$ if and only if $w \in R_+^q$ and

$$\sup_{x \in C_H} \inf_{y \in D_H} \{ \langle x, -s \rangle + \langle y, -t \rangle + H_{z, w}(x, y) \} > -\infty, \quad \forall s \in R^m \quad \forall z \in R_+^p.$$

Now for any fixed $w \in R_+^q$ this condition is easily seen to be equivalent to

$$\exists x \in C_H \text{ such that } \inf_{y \in D_H} \{ \langle y, -t \rangle + H(x, y) + \sum w_j f_j(y) \} > -\infty,$$

that is,

$$t \in \bigcup_{x \in C_H} \text{dom}(H(x, \cdot) + \sum w_j f_j)^*.$$

But for each $w \in R_+^q$ and $x \in C_H$, (16.4) and (16.1) imply that

$$\text{dom}(H(x, \cdot) + \sum w_j f_j)^* = \text{dom} H(x, \cdot)^* + \sum w_j \text{dom} f_j^*,$$

and by Lemma 0.4 we have that

$$\text{dom}_2 H^* = \bigcup_{x \in C_H} \text{dom} H(x, \cdot)^*.$$

It follows that $(t, w) \in \text{dom}_2 L$ is equivalent to

$$w \in R_+^q \text{ and } t \in \bigcup_{x \in C_H} (\text{dom} H(x, \cdot)^* + \sum w_j \text{dom} f_j^*),$$

which is the same as

$$w \in R_+^q \text{ and } t \in \text{dom}_2 H^* + \sum w_j \text{dom} f_j^* .$$

The characterization of $\text{dom}_1 L$ is established similarly, using the other element of $[L]$ described by Theorem 6.8.

We can now describe the "dual objective functions" and the "dual feasible solutions" corresponding to the ordinary saddle program $S(K)$.

THEOREM 6.9. Assume that the program $S(L)$ dual to $S(K)$ is strongly consistent. Then the dual problem of $S(K)$ is well-defined and is given by an equivalence class $[L_0]$ of closed proper convex-concave functions (the "dual objective functions"). The class $[L_0]$ contains both the function

$$(z, w) \rightarrow \begin{cases} \sup_{x \in C_H} \inf_{y \in D_H} \{H(x, y) + \sum z_i g_i(x) + \sum w_j f_j(y)\} & \text{if } z \in R_+^p \text{ and } w \in R_+^q \\ -\infty & \text{if } z \in R_+^p \text{ and } x \notin R_+^q \\ +\infty & \text{if } z \notin R_+^p \end{cases}$$

and the function obtained from this formula by interchanging $\sup_{x \in C_H}$ with

$\inf_{y \in D_H}$. A pair (z, w) is a feasible solution of $S(L)$ (is a "dual feasible

solution") if and only if $(z, w) \in R_+^p \times R_+^q$,

$$0 \in \text{dom}_1 H^* + \sum z_i \text{dom} g_i^* \text{ and } 0 \in \text{dom}_2 H^* + \sum w_j \text{dom} f_j^* ,$$

and the set of such pairs (z, w) is exactly $\text{dom} L_0$.

PROOF. Immediate from Theorem 5.2 applied to $S(L)$, together with Theorems 6.7 and 6.8. Notice that the first characterization of the feasible

solutions of $S(L)$ also follows directly from Theorem 6.8 without the aid of dual strong consistency, once one observes that (z, w) is a feasible solution of $S(L)$ if and only if $(0, z, 0, w) \in \text{dom } L$.

§7. Saddle Programs of Fenchel Type

In this final section we apply the general theory of §5 to another class of minimax problems. Our results extend to minimax theory those obtained by Fenchel [23] and Rockafellar [45, 46, 48] for a certain class of convex optimization problems. These problems enjoy a pleasing symmetry property (not shared by ordinary saddle programs); namely, the dual of such a problem is of the same form. This class includes the minimax problems which Lebedev-Tynjanskii [34] and Tynjanskii [61] considered in an effort to define the dual of a game and also those studied by Rockafellar [47] in connection with double duality theory. The results in [34], [61] and some of those in [60] are improved. Taken together, Theorems 7.3 through 7.8 can be viewed as an extension of the Fenchel-Rockafellar Duality Theorem to saddle functions and minimax problems.

Throughout this section let K be a closed proper concave-convex function on $R^m \times R^n$, L a closed proper convex-concave function on $R^p \times R^q$, $A = A_1 \times A_2$ a linear transformation from $R^m \times R^n$ to $R^p \times R^q$, and define

$$\begin{aligned} X &= \{x \in \text{dom}_1 K \mid A_1 x \in \text{dom}_1 L\} , \\ Y &= \{y \in \text{dom}_2 K \mid A_2 y \in \text{dom}_2 L\} , \\ Z &= \{z \in \text{dom}_1 L^* \mid A_1^* z \in \text{dom}_1 K^*\} , \\ W &= \{w \in \text{dom}_2 L^* \mid A_2^* w \in \text{dom}_2 K^*\} . \end{aligned}$$

Consider the following pair of minimax problems:

(I) Find the saddle points of $K - LA$ with respect to $X \times Y$;

(II) Find the saddle points of $L^* - K^* A^*$ with respect to $Z \times W$.

The problems studied by Rockafellar [47] correspond to the choice $m = p$, $n = q$, A the identity transformation, and L given by

$$L(x, y) = \begin{cases} 0 & \text{if } x \in R_+^m \text{ and } y \in R_+^n \\ +\infty & \text{if } x \notin R_+^m \text{ and } y \in R_+^n \\ -\infty & \text{if } y \notin R_+^n \end{cases}.$$

To apply the results of §5 to these problems, define a function Φ on $(R^p \times R^m) \times (R^q \times R^n)$ by

$$\Phi(u, x, v, y) = \begin{cases} K(x, y) - L(u + A_1 x, v + A_2 y) & \text{if } (u, x) \in \Gamma \text{ and } (v, y) \in \Delta \\ +\infty & \text{if } (u, x) \in \Gamma \text{ and } (v, y) \notin \Delta \\ -\infty & \text{if } (u, x) \notin \Gamma \end{cases},$$

where

$$\Gamma = \{(u, x) \mid x \in \text{dom}_1 K, u + A_1 x \in \text{dom}_1 L\},$$

$$\Delta = \{(v, y) \mid y \in \text{dom}_2 K, v + A_2 y \in \text{dom}_2 L\}.$$

LEMMA 7.1. The function Φ is closed proper concave-convex with domain $\Gamma \times \Delta$, and

$$\text{ri } \Gamma = \{(u, x) \mid x \in \text{ri}(\text{dom}_1 K), u + A_1 x \in \text{ri}(\text{dom}_1 L)\},$$

$$\text{ri } \Delta = \{(v, y) \mid y \in \text{ri}(\text{dom}_2 K), v + A_2 y \in \text{ri}(\text{dom}_2 L)\}.$$

PROOF. Trivially, Γ is convex. Define $\Gamma_x = \{u \mid (u, x) \in \Gamma\}$ for each x . Then Γ_x is empty when $x \notin \text{dom}_1 K$ and equals $\text{dom}_1 L - A_1 x$ when

$x \in \text{dom}_1 K$. Hence (6.8) implies that $(u, x) \in \text{ri } \Gamma$ if and only if $x \in \text{ri}(\text{dom}_1 K)$ and $u \in \text{ri}(\text{dom}_1 L - A_1 x)$. This establishes the formula for $\text{ri } \Gamma$, and the one for $\text{ri } \Delta$ is similar. From these formulas and the fact that K and L are closed and proper, it is not hard to verify (using (34.3)) that Φ has the properties asserted.

According to Lemma 7.1, Φ determines a generalized saddle program $S(\Phi)$ on $R^m \times R^n$ with perturbations in $R^p \times R^q$. The formulas given for $\text{ri } \Gamma$ and $\text{ri } \Delta$ imply that $S(\Phi)$ is strongly consistent if and only if $\text{ri}(\text{dom } L) \cap A \text{ri}(\text{dom } K) \neq \emptyset$. In this event Theorems 1.8, 3.2 and 5.2 imply that $[K - LA]$ is well-defined and gives the primal problem of $S(\Phi)$, which by (36.3) is the same as (I). We say that a generalized saddle program having the form of $S(\Phi)$ is of Fenchel type.

It can be computed as an exercise that the generalized saddle program $S(\Psi)$ dual to $S(\Phi)$ may be given by

$$\Psi(s, z, t, w) = \begin{cases} L^*(z, w) - K^*(s + A_1^* z, t + A_2^* w) & \text{if } (s, z) \in \Pi \text{ and } (t, w) \in \Omega \\ -\infty & \text{if } (s, z) \notin \Pi \text{ and } (t, w) \in \Omega \\ +\infty & \text{if } (t, w) \notin \Omega, \end{cases}$$

where

$$\begin{aligned} \Pi &= \{(s, z) \mid z \in \text{dom}_1 L^*, s + A_1^* z \in \text{dom}_1 K^*\}, \\ \Omega &= \{(t, w) \mid w \in \text{dom}_2 L^*, t + A_2^* w \in \text{dom}_2 K^*\}. \end{aligned}$$

Thus, the dual of a program of Fenchel type is another program of Fenchel type. Hence we have immediately that $S(\Psi)$ is strongly consistent if and

only if $\text{ri}(\text{dom } K^*) \cap A^* \text{ri}(\text{dom } L^*) \neq \emptyset$, and in this event $[L^* - K^* A^*]$ is well-defined and gives the primal problem of $S(\Psi)$ (i.e. the dual problem of $S(\Phi)$). This problem is the same as (II).

With these facts in mind, it is clear that all the results of §5 yield assertions about problems (I) and (II). In the remainder of this section we illustrate some of this.

First we dualize the hypothesis $\text{ri}(\text{dom } K^*) \cap A^* \text{ri}(\text{dom } L^*) \neq \emptyset$. Note that the necessary and sufficient conditions given below are satisfied, for example, when $X \times Y$ is bounded.

LEMMA 7.2. In order that $\text{ri}(\text{dom } K^*) \cap A^* \text{ri}(\text{dom } L^*) \neq \emptyset$, it is necessary and sufficient that

$$(\text{rec}_1 K)(x) \geq (\text{rec}_1 L)(A_1 x) \text{ imply } (\text{rec}_1 K)(-x) \geq (\text{rec}_1 L)(-A_1 x)$$

and

$$(\text{rec}_2 K)(y) \leq (\text{rec}_2 L)(A_2 y) \text{ imply } (\text{rec}_2 K)(-y) \leq (\text{rec}_2 L)(-A_2 y) .$$

PROOF. The lemma will follow from Lemma 5.4, once it is verified that $(\text{rec}_1 \Phi)(0, x) \geq 0$ if and only if $(\text{rec}_1 K)(x) \geq (\text{rec}_1 L)(A_1 x)$ and $(\text{rec}_2 \Phi)(0, y) \leq 0$ if and only if $(\text{rec}_2 K)(y) \leq (\text{rec}_2 L)(A_2 y)$. Only the second equivalence will be checked, as the first is analogous. For each $(u, x) \in \text{ri } \Gamma$, it follows from Lemma 7.1, (9.3) and (9.5) that

$$\text{rec } \Phi(u, x, \cdot, \cdot)(v, y) = \text{rec } K(x, \cdot)(y) - \text{rec } L(u + A_1 x, \cdot)(v + A_2 y) .$$

Hence, $(\text{rec}_2 \Phi)(0, y) \leq 0$ if and only if $\text{rec } K(x, \cdot)(y) \leq \text{rec } L(u + A_1 x, \cdot)(A_2 y)$ holds for each $(u, x) \in \text{ri } \Gamma$. But this latter condition occurs if and only if

$\text{rec } K(x, \cdot)(y) \leq \text{rec } L(u, \cdot)(A_2 y)$ holds for each $x \in \text{ri}(\text{dom}_1 K)$ and $u \in \text{ri}(\text{dom}_1 L)$, which occurs if and only if $(\text{rec}_2 K)(y) \leq (\text{rec}_2 L)(A_2 y)$.

The following result gives "boxing in" inequalities for the lower and upper saddle values pertaining to (I) and (II).

THEOREM 7.3. If $\text{ri}(\text{dom } L) \cap A \text{ri}(\text{dom } K) \neq \emptyset$, then

$$\sup_X \inf_Y (K - LA) \leq \sup_W \inf_Z (L^* - K^* A^*) \leq \inf_Z \sup_W (L^* - K^* A^*) \leq \inf_Y \sup_X (K - LA).$$

Dually, if $\text{ri}(\text{dom } K^*) \cap A^* \text{ri}(\text{dom } L^*) \neq \emptyset$, then

$$\sup_W \inf_Z (L^* - K^* A^*) \leq \sup_X \inf_Y (K - LA) \leq \inf_Y \sup_X (K - LA) \leq \inf_Z \sup_W (L^* - K^* A^*).$$

PROOF. By Corollary 5.5.2, with the aid of (36.4) and (36.3).

It can be shown that the least member of the Lagrangian of $S(\Phi)$ is the function

$$(x, w, y, z) \rightarrow \begin{cases} \underline{K}(x, y) + \underline{L}^*(z, w) - \langle A_1 x, z \rangle - \langle A_2 y, w \rangle & \text{if } (x, w) \in C \text{ and } (y, z) \in \text{cl } D \\ +\infty & \text{if } (x, w) \in C \text{ and } (y, z) \notin \text{cl } D \\ -\infty & \text{if } (x, w) \notin C \end{cases}$$

where $C \times D = (\text{dom}_1 K \times \text{dom}_2 L^*) \times (\text{dom}_2 K \times \text{dom}_1 L^*)$ is the effective domain of the Lagrangian. From this it follows easily by (36.3), (36.4) and (37.4) that two pairs $(x, y) \in R^m \times R^n$ and $(z, w) \in R^p \times R^q$ satisfy the extremality conditions associated with $S(\Phi)$ and $S(\Psi)$ if and only if

$$A(x, y) \in \partial L^*(z, w) \text{ and } A^*(z, w) \in \partial K(x, y). \quad (**)$$

To state the remaining results we need the following definitions. The optimal value of (I) is the saddle value of $K - LA$ with respect to $X \times Y$

(when this saddle value exists). An optimal solution of (I) is a saddle point of $K - LA$ with respect to $X \times Y$. It is convenient to say that an optimal solution of (I) is stable if and only if it is a stable optimal solution of $S(\Phi)$. Similar definitions are used for (II).

THEOREM 7.4. A pair (x, y) is a stable optimal solution of (I) if and only if there exists a pair (z, w) such that $(**)$ holds. Dually, a pair (z, w) is a stable optimal solution of (II) if and only if there exists a pair (x, y) such that $(**)$ holds.

PROOF. By Theorem 5.13.

THEOREM 7.5. Problem (I) has a stable optimal solution if and only if problem (II) does, in which case the optimal values in (I) and (II) are equal.

PROOF. By Theorem 5.12.

THEOREM 7.6. If $\text{ri}(\text{dom } L) \cap A \text{ ri}(\text{dom } K) \neq \emptyset$, then every optimal solution of (I) is stable. Dually, if $\text{ri}(\text{dom } K^*) \cap A^* \text{ ri}(\text{dom } L^*) \neq \emptyset$, then every optimal solution of (II) is stable.

PROOF. By Theorem 5.15.

To go along with these results there is the following general existence theorem. (We omit the dual version.)

THEOREM 7.7. Assume that $\text{ri}(\text{dom } L) \cap A \text{ ri}(\text{dom } K) \neq \emptyset$ and that either $X \times Y$ is bounded or (more generally) that the following two conditions are satisfied:

(a) If $\text{rec } K(x, \cdot)(y) \leq \text{rec } L(A_1 x, \cdot)(A_2 y)$ for every $x \in \text{ri } X$, then $\text{rec } K(x, \cdot)(-y) \leq \text{rec } L(A_1 x, \cdot)(-A_2 y)$ for every $x \in \text{ri } X$.

(b) If $\text{rec } K(\cdot, y)(x) \geq \text{rec } L(\cdot, A_2 y)(A_1 x)$ for every $y \in \text{ri } Y$, then
 $\text{rec } K(\cdot, y)(-x) \geq \text{rec } L(\cdot, A_2 y)(-A_1 x)$ for every $y \in \text{ri } Y$.

Then there exists an optimal solution of (I).

PROOF. Since $[\Phi_0] = [K - LA]$, the theorem will follow immediately from Corollary 5.2.1 once it is checked that $\text{rec}_2(K - LA)(y) \leq 0$ if and only if $\text{rec } K(x, \cdot)(y) \leq \text{rec } L(A_1 x, \cdot)(A_2 y)$ for every $x \in \text{ri } X$, and that $\text{rec}_1(K - LA)(x) \geq 0$ if and only if $\text{rec } K(\cdot, y)(x) \geq \text{rec } L(\cdot, A_2 y)(A_1 x)$ for every $y \in \text{ri } Y$. We show only the first equivalence, as the second is similar. By Theorems 3.2 and 1.8, $[K - LA]$ has effective domain $X \times Y$ and contains the function H given by

$$H(x, y) = \begin{cases} K(x, y) - LA(x, y) & \text{if } x \in X \text{ and } y \in Y \\ +\infty & \text{if } x \in X \text{ and } y \notin Y \\ -\infty & \text{if } x \notin X \end{cases}$$

From this together with (9.3) and (9.5), it follows that

$$\text{rec } H(x, \cdot)(y) = \text{rec } K(x, \cdot)(y) - \text{rec } L(A_1 x, \cdot)(A_2 y)$$

for every $x \in \text{ri } X$. Since $\text{rec}_2(K - LA)(y) = \sup\{\text{rec } H(x, \cdot)(y) \mid x \in \text{ri } X\}$, the equivalence follows.

The preceding results can be collected together quite concisely in the presence of strong consistency and boundedness assumptions.

THEOREM 7.8. Assume either that $\text{ri}(\text{dom } L) \cap A \text{ ri}(\text{dom } K) \neq \emptyset$ and $X \times Y$ is bounded or dually that $\text{ri}(\text{dom } K) \cap A^* \text{ ri}(\text{dom } L^*) \neq \emptyset$ and $Z \times W$ is

bounded. Then both (I) and (II) have optimal solutions and the optimal values in (I) and (II) are equal. Moreover, these optimal solutions are stable, and they are precisely the pairs (x, y) and (z, w) for which (**)
holds.

Appendix: Polyhedral Refinements

Recall from §0 that a saddle function is polyhedral if and only if it is closed and either its concave or its convex parent is polyhedral. It has been noted that each of the operations developed in the paper preserve this property of polyhedralness. Much more can be said, however. In fact, nearly all the results in the paper admit refinements when some or all of the saddle functions involved are polyhedral. These refinements are, for the most part, not hard to establish. One simply needs to make slight, systematic modifications of the existing proofs. These revisions rest ultimately on just a few additional results which will be presented below, plus two basic "principles."

The first of these principles can be stated as follows. If certain conclusions can be deduced from a condition of the form $(C \times D) \cap \text{ri}(\text{dom } K) \neq \emptyset$ when K is a closed proper saddle function and C and D are convex sets, then the same conclusions (at least) can be deduced from the weaker condition $(C \times D) \cap \text{dom } K \neq \emptyset$ when K is actually polyhedral. "Overlapping" conditions of this kind are generally to ensure that some operation involving $[K]$ can be performed to yield a well-defined new equivalence class. The principle just says that less overlapping is required for well-definedness when $[K]$ is polyhedral. The implementation of this principle rests essentially on the theorem below, which is usually used in place of (34.3) in proving the polyhedral refinements.

THEOREM A.1. Let K be a polyhedral proper concave-convex function on $R^m \times R^n$. Then (i) the sets $\text{dom}_1 K$ and $\text{dom}_2 K$ are polyhedral convex

(hence closed), (ii) the elements of $[K]$ agree everywhere on $(R^m \times \text{dom}_2 K) \cup (\text{dom}_1 K \times R^n)$, (iii) for each $x \in \text{dom}_1 K$ the function $K(x, \cdot)$ is polyhedral proper convex (hence closed) with effective domain $\text{dom}_2 K$, and (iv) for each $y \in \text{dom}_2 K$ the function $K(\cdot, y)$ is polyhedral proper concave (hence closed) with effective domain $\text{dom}_1 K$.

PROOF. Assertion (i) follows from the representations of $\text{dom}_j K$ given in Theorem 0.1(a) and the fact that linear transformations preserve polyhedral convex sets (19.3). Assertion (ii) is just a restatement of (33.2.2). Now suppose $x \in \text{dom}_1 K$. Then (34.3) implies $K(x, \cdot)$ is proper convex with its effective domain between $\text{dom}_2 K$ and $\text{cl}(\text{dom}_2 K)$. But since $\text{dom}_2 K$ is polyhedral convex, it is closed (19.1), and hence $\text{dom} K(x, \cdot) = \text{dom}_2 K$. Also, from (ii) and Theorem 0.1(b) it follows that $K(x, \cdot) = \underline{K}(x, \cdot) = f(x, \cdot)^*$. Hence (19.3.1) and (19.2) imply that $K(x, \cdot)$ is polyhedral. This establishes (iii), and (iv) is proved similarly.

In order to discuss the second "principle" we need two more polyhedral results. Recall that Lemma 0.7 was the tool which enabled us to dualize various conditions throughout the paper (e.g. Lemmas 1.9, 3.5, 5.4, 6.6 and 7.2). This next lemma can be used similarly to dualize the polyhedral versions of the same conditions.

LEMMA A.2. Let K be a polyhedral proper concave-convex function on $R^m \times R^n$. If L is a subspace of R^m , then $L \cap \text{dom}_1 K^* \neq \emptyset$ if and only if $(\text{rec}_1 K)(x) \leq 0$ for every $x \in L^\perp$. Similarly, if L is a subspace of R^n , then $L \cap \text{dom}_2 K^* \neq \emptyset$ if and only if $(\text{rec}_2 K)(y) \geq 0$ for every $y \in L^\perp$.

PROOF. We prove only the second equivalence, as the first is similar. Write $\text{dom}_2 K^* = D^*$. By (20.2), $L \cap D^* = \emptyset$ if and only if there exist $y \in R^n$ and $\alpha \in R$ such that

$$\sup_{D^*} \langle \cdot, y \rangle \leq \alpha \leq \inf_L \langle \cdot, y \rangle \text{ and } \alpha < \sup_L \langle \cdot, y \rangle.$$

But the latter condition occurs if and only if there exists $y \in R^n$ such that $\sup_{D^*} \langle \cdot, y \rangle \leq \inf_L \langle \cdot, y \rangle$ and $\sup_{D^*} \langle \cdot, y \rangle < \sup_L \langle \cdot, y \rangle$, i.e. if and only if there exists $y \in L^\perp$ such that $\sup_{D^*} \langle \cdot, y \rangle < 0$. Since $\sup_{D^*} \langle \cdot, y \rangle = (\text{rec}_2 K)(y)$ by Theorem 0.3, this finishes the proof.

Also needed is the fact that the subdifferential mapping is better behaved when K is polyhedral, as explained below.

THEOREM A. 3. Let K be a polyhedral proper concave-convex function on $R^m \times R^n$. Then $\text{dom } \partial K = \text{dom } K$, and $\partial K(x, y)$ is a product of polyhedral convex sets for each $(x, y) \in \text{dom } K$.

PROOF. Since K is closed and proper, (37.4) implies that $\text{dom } \partial K \subset \text{dom } K$. On the other hand, let $(x, y) \in \text{dom } K$. Then Theorem A.1 (iii) and (23.10) imply that $\partial K(x, \cdot)(y)$ is nonempty polyhedral convex, and similarly Theorem A.1 (iv) and (23.10) imply $\partial K(\cdot, y)(x)$ is nonempty. But $\partial K(x, y) = \partial K(\cdot, y)(x) \times \partial K(x, \cdot)(y)$.

We can now state the second general "principle" involved in the polyhedral refinements. It concerns results which essentially assert the existence of a saddle point for some K . Such results are generally based on the fact (cf. (37.5.3)) that the condition $(0, 0) \in \text{ri}(\text{dom } K^*)$ is sufficient for a closed

proper K to have a saddle point, and that (by Lemma 0.5) this condition can be written dually as the pair of conditions

$$\begin{aligned}(\text{rec}_1 K)(x) \geq 0 &\implies (\text{rec}_1 K)(-x) \geq 0, \\ (\text{rec}_2 K)(y) \leq 0 &\implies (\text{rec}_2 K)(-y) \leq 0.\end{aligned}$$

When K is actually polyhedral, Theorem A.3 and (37.5.3) imply that the condition $(0,0) \in \text{dom } K^*$ is both sufficient and necessary for K to have a saddle point, and by Lemma A.2 this condition can be written dually as the pair of conditions

$$\begin{aligned}(\text{rec}_1 K)(x) &\leq 0, \quad \forall x \\ (\text{rec}_2 K)(y) &\geq 0, \quad \forall y.\end{aligned}$$

The second principle, then, says that when K is polyhedral and proper the former (sufficient) conditions on the recession functions of K can be relaxed to the latter (necessary and sufficient) conditions in results which assert the existence of a saddle point for K .

Examples of results whose polyhedral refinements involve applying the first principle are Theorems 1.2, 3.6, 5.2, 1.8, 3.10, 5.5, and 3.2. Results whose refinements entail applying both principles include Theorems 2.4, 2.5, 3.11, and 5.6, Corollary 5.2.1, and Theorems 6.3 and 7.7. The refinements of Theorems 2.4 and 2.5 are the most involved.

In some places it is helpful to have the flexibility furnished by the various representations of $\text{rec}_j K$ given in the next result.

THEOREM A.4. Let K be a closed proper concave-convex function on $\mathbb{R}^m \times \mathbb{R}^n$. If C is any set satisfying $\text{ri}(\text{dom}_1 K) \subset C \subset \text{dom}_1 K$, then

$$(\text{rec}_2 K)(y) = \sup_{x \in C} \{ \text{rec } \underline{K}(x, \cdot)(y) \} .$$

Similarly, if D is any set satisfying $\text{ri}(\text{dom}_2 K) \subset D \subset \text{dom}_2 K$, then

$$(\text{rec}_1 K)(x) = \inf_{y \in D} \{ \text{rec } \bar{K}(\cdot, y)(x) \} .$$

PROOF. We prove only the first assertion, as the second is similar.

Write $\text{dom}_2 K^* = D^*$. By Theorem 0.3 we know that $\text{rec}_2 K = \delta^*(\cdot | D^*)$. Now from the formulas given for D^* and $\text{ri } D^*$ in the proof of (37.2), it is clear that

$$\text{ri } D^* \subset \bigcup_{x \in C} \text{dom } f(x, \cdot) \subset D^* ,$$

where f is the convex parent of K . Hence, much as in the proof of (37.2),

we can write $\delta^*(y | D^*) = \sup \{ \langle y^*, y \rangle | y \in \text{dom } f(x, \cdot), x \in C \} =$

$\sup_{x \in C} \{ \delta^*(y | \text{dom } f(x, \cdot)) \} = \sup_{x \in C} \{ \text{rec } f(x, \cdot)^*(y) \} = \sup_{x \in C} \{ \text{rec } \underline{K}(x, \cdot)(y) \}$, as asserted.

In proving polyhedral refinements there are naturally places where polyhedral refinements of various results from convex function theory are needed. With few exceptions, these supplementary facts are mentioned explicitly in [48] (e.g. (19.2)). In particular, the important tool (16.3) has a polyhedral refinement which is indicated in [48, p. 144]. We include a proof for completeness.

LEMMA A.5. Let f be a polyhedral convex function on R^n , and let $A: R^m \rightarrow R^n$ be a linear transformation. Assume $\text{range } A \cap \text{dom } f \neq \emptyset$, or equivalently, assume that $(\text{rec } f^*)(x^*) \geq 0$ whenever $A^* x^* = 0$. Then $A^* f^*$ is proper and $(fA)^* = A^* f^*$. Moreover, for each y^* the infimum in the definition $(A^* f^*)(y^*) = \inf \{ f^*(x^*) | A^* x^* = y^* \}$ either is attained or is $+\infty$ vacuously.

PROOF. The equivalence of the two conditions in the hypothesis can be deduced from the polyhedral refinement of (16.2), whose proof is just like that of Lemma A.2. By (19.3.1), fA and $A^{**}f^{**}$ are polyhedral convex; moreover, the infimum in the definition of $A^{**}f^{**}$, if finite, is attained. Observe that a polyhedral proper convex function is necessarily closed (19.1.2). Then (16.3) implies $(fA)^* = \text{cl}(A^{**}f^{**})$. Since $\text{range } A \cap \text{dom } f \neq \emptyset$ implies fA is proper, we can conclude that $\text{cl}(A^{**}f^{**})$, and hence $A^{**}f^{**}$, is proper. Finally, the above observation implies that $A^{**}f^{**}$ is closed. This completes the proof.

With the aid of the two "principles" and the additional results just given, one can systematically carry out polyhedral refinements for almost all of the results in the paper except for those in §4 and for the "mixed" cases of §§3, 6 and 7. The results of §4 plainly do not admit polyhedral refinements. It appears that also the mixed case of §6 (i.e. when some of $H, g_1, \dots, g_p, f_1, \dots, f_q$ are polyhedral) does not admit polyhedral refinements. The troublesome spot is in proving a version of Theorem 5.15, i.e. the existence of Lagrange multipliers. However the "mixed" case of §7 follows from that of §3, which we now discuss.

In the event that K_1, \dots, K_s are all polyhedral and proper, the refinements for §3 can of course be derived directly from those of §§1 and 2. But in the mixed case, for example when only K_1, \dots, K_t are polyhedral for some $1 \leq t < s$, a different line of argument is needed. To begin with, one follows the first "principle" and proves a version of Theorem 3.2 by replacing the hypothesis (*) with the hypothesis

$$\text{dom } K_1 \cap \dots \cap \text{dom } K_t \cap \text{ri}(\text{dom } K_{t+1}) \cap \dots \cap \text{ri}(\text{dom } K_s) \neq \emptyset .$$

Then using this hypothesis one must essentially retrace the outline of §§1 and 2, appealing to (20.1) (resp. (23.8)) everywhere the original proofs appeal to (16.3) (resp. (23.9)). In carrying out this program notice that the important tool (6.5) no longer applies to allow full use of the relaxed intersection hypothesis. Fortunately, though, we have the following refinement of (6.5), which when used in conjunction with (6.5) makes a perfectly satisfactory substitute.

LEMMA A.6. Let C_1 and C_2 be convex sets in R^n satisfying
 $C_1 \cap \text{ri } C_2 \neq \emptyset$. Then

$$\text{ri}(C_1 \cap C_2) \subset C_1 \cap \text{ri } C_2$$

and

$$C_1 \cap \text{cl } C_2 \subset \text{cl}(C_1 \cap C_2) .$$

PROOF. Suppose we know that $\text{ri } C_0 \cap \text{ri } C_2 \neq \emptyset$ where C_0 is the intersection of C_1 with the affine hull of C_2 . Noting that $\text{ri } C_2$ and $\text{cl } C_2$ have the same affine hull as C_2 (6.2), we can apply (6.5) to conclude that $\text{ri}(C_1 \cap C_2) = \text{ri}(C_0 \cap C_2) = \text{ri } C_0 \cap \text{ri } C_2 \subset C_0 \cap \text{ri } C_2 = C_1 \cap \text{ri } C_2$ and $C_1 \cap \text{cl } C_2 = C_0 \cap \text{cl } C_2 \subset \text{cl } C_0 \cap \text{cl } C_2 = \text{cl}(C_0 \cap C_2) = \text{cl}(C_1 \cap C_2)$. Thus we are done once we show that $\text{ri } C_0$ meets $\text{ri } C_2$. Assume to the contrary that $\text{ri } C_0 \cap \text{ri } C_2 = \emptyset$. Since C_0 and C_2 are nonempty by the hypothesis, (11.3) implies that there exists a hyperplane H separating C_0 and C_2 properly. For definiteness let $C_0 \subset H \cup A$

and $C_2 \subset H \cup B$, where A and B denote the two open halfspaces determined by H . Again using the hypothesis, pick some point $z \in C_1 \cap \text{ri } C_2$. Clearly $z \in H$. Suppose there existed some point $a \in C_0 \cap A$. Since $z \in \text{ri } C_2$ and a belongs to the affine hull of C_2 , there would exist λ such that $0 < \lambda < 1$ and $x = (1 - \lambda)z + \lambda a \in C_2$. But (6.1) together with $z \in H$ and $a \in A$ would imply $x \in A$. Since this would contradict $C_2 \cap A = \emptyset$, we conclude that $C_0 \subset H$. On the other hand, suppose there existed some point $b \in C_2 \cap B$. Since $z \in \text{ri } C_2$ and $b \in C_2$, (6.4) would imply the existence of some $\mu > 1$ such that $y = (1 - \mu)b + \mu z \in C_2$. But $z \in H$, $b \in B$ and $\mu > 1$ would imply that $y \in A$. Since this would contradict $C_2 \cap A = \emptyset$, we conclude that $C_2 \subset H$. Therefore $C_0 \cup C_2 \subset H$. But now this contradicts the properness of the separation. Since we are led to a contradiction in any case, the original assumption $\text{ri } C_0 \cap \text{ri } C_2 = \emptyset$ must be rejected and the proof is complete.

Finally, notice also for the mixed case of §3 that alternative representations of $\text{rec}_j(K_1 + \dots + K_s)$ are helpful. These are obtained from Lemma A.6, (6.5) and Theorem A.4 via the choices $C = C_1 \cap \dots \cap C_t \cap \text{ri } C_{t+1} \cap \dots \cap \text{ri } C_s$ and $D = D_1 \cap \dots \cap D_t \cap \text{ri } D_{t+1} \cap \dots \cap \text{ri } D_s$, where $C_1 \times D_1 = \text{dom } K_1$.

REFERENCES

- [1] K. J. Arrow and L. Hurwicz, "Reduction of constrained maxima to saddle-point problems," in Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability, J. Neyman, ed., University of California Press, Berkeley, Calif., 1956, vol. V, 1-20.
- [2] K. J. Arrow, L. Hurwicz, and H. Uzawa, Studies in Linear and Non-linear Programming, Stanford University Press, Stanford, Calif., 1958.
- [3] A. Auslender, "Recherche des points de selle d'une fonction," in Symposium on Optimization, A. V. Balakrishnan, M. Contensou, B. F. de Veubeke, P. Kree, J. L. Lions, and N. N. Moiseev, eds., Springer-Verlag, New York, 1970, 37-52.
- [4] C. Berge, "Sur une convexité régulière non linéaire et ses applications à la théorie des jeux," Bull. Soc. Math. France 82(1954), 301-315.
- [5] J. Bram, "The Lagrange multiplier theorem for max-min with several constraints," SIAM J. Appl. Math. 14(1966), 665-667.
- [6] G. W. Brown and J. von Neumann, "Solutions of games by differential equations," in Contributions to the Theory of Games, Vol. I., Annals of Mathematics Studies, No. 24, Princeton University Press, Princeton, N. J., 1950, 73-79.
- [7] Ju. V. Čaikovskiĭ, "On a continuous process of matrix game solution," Soviet Math. Dokl. 12 (1971), 1245-1247.
- [8] A. Charnes, "Constrained games and linear programming," Proc. Natl. Acad. Sci. U. S. 39 (1953), 639-641.

- [9] Y. Cherruault and P. Loridan, "Méthodes pour la recherche de points de selle," C. R. Acad. Sci. Paris 273 (1971), 171-174.
- [10] J. M. Danskin, "Fictitious play for continuous games," Naval Res. Logist. Quart. 1 (1954), 313-320.
- [11] _____, The Theory of Max-Min, Springer-Verlag, New York, 1967.
- [12] _____, "An algorithm for the solution of concave-convex games," in Computing Methods in Optimization Problems -2, L. A. Zadeh, L. W. Neustadt, and A. V. Balakrishnan, eds., Academic Press, New York, 1969, 67-76.
- [13] G. B. Dantzig, "A proof of the equivalence of the programming problem and the game problem," in Activity Analysis of Production and Allocation, T. C. Koopmans, ed., Wiley, New York, 1951, 330-335.
- [14] G. B. Dantzig, E. Eisenberg, and R. W. Cottle, "Symmetric dual nonlinear programs," Pacific J. Math. 15 (1965), 809-812.
- [15] V. F. Demyanov, "Successive approximations for finding saddle points," Soviet Math. Dokl. 8 (1967), 1350-1353.
- [16] _____, "Algorithms for some minimax problems," J. Comput. System Sci. 2 (1968), 342-380.
- [17] _____, "Finding saddle points on polyhedra," Soviet Math. Dokl. 11 (1970), 558-561.
- [18] _____, "On finding saddle points on polyhedrons," Vestnik Leningrad. Univ. 3 (1971), 42-48 (Russian).

- [19] J. E. Falk, "Lagrange multipliers and nonlinear programming,"
J. Math. Anal. Appl. 19 (1967), 141-159.
- [20] K. Fan, "Fixed-point and minimax theorems in locally convex topological linear spaces," Proc. Natl. Acad. Sci. U. S. 38 (1952), 121-126.
- [21] _____, "Minimax theorems," Proc. Natl. Acad. Sci. U. S. 39 (1953), 42-47.
- [22] W. Fenchel, "On conjugate convex functions," Canad. J. Math. 1 (1949), 73-77.
- [23] _____, Convex Cones, Sets and Functions, mimeographed lecture notes, Princeton University, Princeton, N. J., 1951.
- [24] D. Gale, H. W. Kuhn, and A. W. Tucker, "Linear programming and the theory of games," in Activity Analysis of Production and Allocation, T. C. Koopmans, ed., Wiley, New York, 1951, 317-329.
- [25] A. Ghouila-Houri, "Le théorème minimax de SION," in Theory of Games, Techniques and Applications, A. Mensch, ed., American Elsevier, New York, 1966, 123-129.
- [26] I. L. Glicksberg, "A further generalization of the Kakutani fixed point theorem, with application to Nash equilibrium points," Proc. Amer. Math. Soc. 3 (1952), 170-174.
- [27] N. I. Gratchev and Y. G. Evtushenko, "The numerical method of minimax problems solving," Ž. Vyčisl. Mat. i Mat. Fiz. 11 (1971), 313-327 (Russian). English translation to appear in U. S. S. R. Comput. Math. and Math. Phys.

- [28] A. D. Ioffe and V. M. Tikhomirov, "Duality of convex functions and extremum problems," Russian Math. Surveys 23(1968), no. 6, 53-124.
- [29] S. Kakutani, "A generalization of Brouwer's fixed point theorem," Duke Math. J. 8 (1941), 457-459.
- [30] H. Kneser, "Sur un théorème fondamental de la théorie des jeux," C. R. Acad. Sci. Paris 234 (1952), 2418-2420.
- [31] H. W. Kuhn and G. P. Szegö (editors), Differential Games and Related Topics, American Elsevier, New York, 1971.
- [32] H. W. Kuhn and A. W. Tucker, "Nonlinear programming," in Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability, University of California Press, Berkeley, Calif., 1951, 481-492.
- [33] L. S. Lasdon, Optimization Theory for Large Systems, MacMillan, New York, 1970.
- [34] V. N. Lebedev and H. T. Tynjanskii, "Duality theory for concave-convex games," Soviet Math. Dokl. 8 (1967), 752-756.
- [35] O. L. Mangasarian and J. Ponstein, "Minimax and duality in non-linear programming," J. Math. Anal. Appl. 11 (1965), 504-518.
- [36] J. -J. Moreau, "Sur la fonction polaire d'une fonction semicontinue supérieurement," C. R. Acad. Sci. Paris 258 (1964), 1128-1131.
- [37] _____, "Théorèmes 'inf-sup'," C. R. Acad. Sci. Paris 258 (1964), 2720-2722.

- [38] J. von Neumann, "Zur Theorie der Gesellschaftsspiele," Math. Ann. 100(1928), 295-320.
- [39] H. Nikaidō, "On von Neumann's minimax theorem," Pacific J. Math. 4 (1954), 65-72.
- [40] B. N. Pshenichnyi, "The dual method in extremum problems. I, II," Cybernetics I (1965), no. 3, 91-99, no. 4, 72-79.
- [41] J. Robinson, "An iterative method of solving a game," Ann. of Math. (2) 54 (1951), 296-301.
- [42] R. T. Rockafellar, Convex Functions and Dual Extremum Problems, thesis, Harvard University, Cambridge, Mass., 1963.
- [43] _____, "Minimax theorems and conjugate saddle functions," Math. Scand. 14(1964), 151-173.
- [44] _____, "Helly's theorem and minima of convex functions," Duke Math. J. 32 (1965), 381-397.
- [45] _____, "An extension of Fenchel's duality theorem for convex functions," Duke Math. J. 33 (1966), 81-89.
- [46] _____, "Duality and stability in extremum problems involving convex functions," Pacific J. Math. 21(1967), 167-187.
- [47] _____, "A general correspondence between dual minimax problems and convex programs," Pacific J. Math. 25(1968), 597-611.
- [48] _____, Convex Analysis, Princeton University Press, Princeton, N. J., 1970.

- [49] _____, "Monotone operators associated with saddle-functions and minimax problems," in Nonlinear Functional Analysis, F. E. Browder, ed., Proceedings of Symposia in Pure Mathematics, Vol. XVIII, Part 1, American Mathematical Society, Providence, R. I., 1970, 241-250.
- [50] _____, "Conjugate convex functions in optimal control and the calculus of variations," J. Math. Anal. Appl. 32 (1970), 174-222.
- [51] _____, "Generalized Hamiltonian equations for convex problems of Lagrange," Pacific J. Math. 33 (1970), 411-427.
- [52] _____, "Saddle-points and convex analysis," in Differential Games and Related Topics, H. W. Kuhn and G. P. Szegö, eds., American Elsevier, New York, 1971, 109-127.
- [53] M. Shiffman, "On the equality $\min \max = \max \min$, and the theory of games," RAND Report RM-243, 1949.
- [54] M. Slon, "On general minimax theorems," Pacific J. Math. 8 (1958), 171-176.
- [55] M. Slater, "Lagrange multipliers revisited: a contribution to non-linear programming," Cowles Commission Discussion Paper, Math. 403 (1950).
- [56] J. Stoer, "Duality in nonlinear programming and the minimax theorem," Numer. Math. 5 (1963), 371-379.
- [57] _____, "Über einen Dualitätssatz der nichtlinearen Programmierung," Numer. Math. 6 (1964), 55-58.

- [58] J. Stoer and C. Witzgall, Convexity and Optimization in Finite Dimensions I, Springer-Verlag, New York, 1970.
- [59] R. Trémolières, "Algorithmes pour la résolution de problèmes posés sous la forme 'min. max. $f(x, y)$.' Cas non linéaire," report no. Inf. 6909, Institut de Recherche d'Informatique et d'Automatique, Rocquencourt, May, 1969.
- [60] N. T. Tynjanskii, "Conjugate concave-convex functions in linear topological spaces and their saddle-points," Mat. Sb. 7 (1969), 503-531.
- [61] _____, "General concave-convex games," Soviet Math. Dokl. 10 (1969), 91-94.
- [62] J. Ville, "Sur la théorie générale des jeux où intervient l'habilité des joueurs," in Traité du Calcul des Probabilités et de ses Applications 4 (1938), no. 2, 105-113.
- [63] A. Wald, "Generalization of a theorem by v. Neumann concerning zero sum two person games," Ann. of Math. 46 (1945), 281-286.